

DIRECT-SEQUENCE CODE-DIVISION MULTIPLE ACCESS SYSTEMS:
NEAR-FAR RESISTANT PARAMETER ESTIMATION AND DETECTION

By
ERIK G. STRÖM

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1994

To my family, Karl and Ann-Britt Ström, Johan Ström, and Karin Ström, and to my love, Maria Lundqvist.

ACKNOWLEDGMENTS

I would like to express my gratitude to my mentors, Dr. Scott L. Miller and Dr. Fred J. Taylor. Without their guidance, support, and encouragement, I would never have been able to complete this dissertation. I thank the members of my Ph.D. committee, Drs. Jose Principe, Kermit Sigmon and Donald Childers, for many fruitful discussions.

I am grateful for the financial support from Motorola Inc., Plantation, FL, and from the Swedish National Board for Industrial and Technical Development.

The majority of the parameter estimation work presented in Chapters 4 and 7 was done in collaboration with Stefan Parkvall and Dr. Björn E. Ottersten of the Royal Institute of Technology, Stockholm, Sweden.

I would also like to mention a few people who made my stay in Gainesville and at the University of Florida such a wonderful experience: Drs. Tex and Joan Lindgren, Jan-Erik Lindgren, Fred and Anna Stengård, David and Tara Petty, the Robbins, Viktoria Öberg, Andrew Martin, Tang Forbes, Monica Murphy, Jon and Jane Mellott, John and Ann Ruckstuhl, Ahmad Ansari, Jeremy Smith, Iztok Koren, and Steve Spurrier.

I will end this acknowledgment with two poems from *Grooks* by Piet Hein, which I recite to myself when things are not working out as I wish they would:

—T.T.T.—

Put up in a place
where it's easy to see
the cryptic admonishment

T.T.T.

When you feel how depressingly
slowly you climb,
it's well to remember that
Things Take Time.

—The Road to Wisdom—

The road to wisdom?—Well, it's plain
and simple to express

Err
and err
and err again
but less
and less
and less

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
KEY TO ABBREVIATIONS	viii
KEY TO SYMBOLS	x
ABSTRACT	xvii
CHAPTERS	
1 INTRODUCTION	1
1.1 Notation and Conventions	1
1.2 Introduction to Multiple-Access Communication Systems	2
1.3 Dissertation Outline	9
2 SYSTEM MODEL FOR THE AWGN CHANNEL	10
2.1 Continuous-Time System Model	10
2.2 Discrete-Time Complex Vector System Model	11
2.3 Properties of the Correlation Matrix	14
3 RECEIVERS FOR THE AWGN CHANNEL	18
3.1 Introduction to DS-CDMA Receivers	18
3.2 Generic Linear Receiver	21
3.2.1 Probability of Error	22
3.2.2 The Near-Far Problem	25
3.3 Standard Receiver	27
3.4 Single-Bit, Single-User Decorrelating Receiver	28
3.5 Minimum Mean-Squared Error Receiver	29
3.5.1 Gaussian Approximation Error Probability for the MMSE Receiver	30
3.5.2 Asymptotic Properties of the MMSE Receiver	32
3.6 Dimension Reduction	36
3.6.1 Cyclically Shifted Filter Bank	41
3.6.2 Gamma and Modified Gamma	42
3.6.3 Symmetric Dimension Reduction	44

3.6.4	Eigenvector Identification	45
3.6.5	Numerical Results	46
3.7	Linear Conjugate Linear MMSE Receiver	49
3.7.1	Linear Conjugate Linear Filtering	49
3.7.2	MMSE Receiver with A Priori Phase Knowledge	52
3.7.3	Simulation of LCL-MMSE and MMSE Receivers.	54
3.8	Conclusions	56
4	PARAMETER ESTIMATION FOR THE AWGN CHANNEL	59
4.1	Cramér-Rao Bound	60
4.2	Estimation Algorithms	64
4.2.1	Maximum-Likelihood	64
4.2.2	Approximative Maximum Likelihood	65
4.2.3	MUSIC	66
4.2.4	Sliding Correlator	68
4.2.5	Numerical Results	69
4.3	Dimension Reduction	73
4.3.1	Optimum Dimension Reduction	73
4.3.2	Numerical Results	76
4.4	Conclusions	77
5	SENSITIVITY ANALYSIS OF THE DECORRELATING RECEIVER	79
5.1	Introduction	79
5.2	System Model	80
5.3	Analysis of Linear Receivers	83
5.4	Analysis of the Decorrelating Receiver	85
5.5	Numerical Results	85
5.6	Conclusions	88
6	SYSTEM MODEL FOR FADING CHANNELS	89
6.1	Continuous-Time Channel Model	89
6.2	Discrete-Time Complex Vector System Model	90
7	PARAMETER ESTIMATION FOR FADING CHANNELS	93
7.1	Cramér-Rao Bound	94
7.2	Estimation Algorithms	98
7.2.1	Maximum-Likelihood	98
7.2.2	Approximative Maximum-Likelihood	99
7.2.3	MUSIC	99
7.2.4	Numerical Results	100
7.3	Dimension Reduction	103
7.4	Conclusions	105

8	SUMMARY AND CONCLUSIONS	107
8.1	Summary	107
8.2	Contributions	109
8.3	Future Work	110
APPENDICES		
A	DERIVATION OF THE CRAMÉR-RAO BOUND	112
B	TYPOGRAPHICAL NOTES	132
REFERENCES		133
BIOGRAPHICAL SKETCH		139

KEY TO ABBREVIATIONS

ADC: American digital standard
AML: approximative maximum likelihood
AWGN: additive white Gaussian noise
BPSK: bi-polar phase shift keying
CDMA: code-division multiple access
CRB: Cramér-Rao bound
CSFB: cyclically shifted filter bank
DS-CDMA: direct-sequence code-division multiple access
EGID: eigenvector identification
EM: expectation-maximization
FDMA: frequency-division multiple-access
FH-CDMA: frequency-hopping code-division multiple access
FIR: finite impulse response
IEEE: Institute of Electrical and Electronics Engineers
IQ: in-phase quadrature-phase
JTIDS: joint tactical distribution system
LCL-MMSE: linear-conjugate-linear MMSE
LMS: least mean squared
MAI: multiple-access interference
ML: maximum likelihood
MMSE: minimum mean-squared error

MRASE: multiple-access rapid acquisition by sequential estimation

MSE: mean-squared error

MUSIC: multiple signal classification

PSD: power spectral density

RASE: rapid acquisition by sequential estimation

RLS: recursive least-squares

SDR: symmetric dimension reduction

SNR: signal-to-noise ratio

SVD: singular value decomposition

TDMA: time-division multiple-access

KEY TO SYMBOLS

Below follows a list of the main symbols used in this dissertation. However, no symbols from Chapter 5 are listed here since the notation in that chapter is quite different from—and sometimes in conflict with—the notation used in the other chapters. If not otherwise indicated, the symbols represent real or integer numbers. As might be suspected, the number displayed at the end of the dotted line is the number of the page upon which the symbol first appeared.

$\mathbf{A} \in \mathbb{R}^{QN \times 2K}$, matrix of received vectors (AWGN)	13
$\mathbf{A} \in \mathbb{R}^{QN \times 2R}$, matrix of received vectors (fading channel)	92
$\mathbf{A}_k \in \mathbb{R}^{QN \times 2R_k}$, matrix of k th user's received vectors (fading)	91
$\mathbf{a}_{2k-1}, \mathbf{a}_{2k} \in \mathbb{R}^{QN}$, k th user's received vectors	13
$\alpha_{k,r}(t)$, fading process of k th user's r th ray	89
$\mathbf{A}' \in \mathbb{R}^{QN \times 2(K-1)}$, matrix of received vectors due to interfering users (AWGN) ...	24
$\mathbf{A}_V = \mathbf{V}^* \mathbf{A} \in \mathbb{C}^{L \times 2R}$, reduced dimension $\mathbf{A}$ matrix (fading)	104
$\mathbf{A}_V \in \mathbb{C}^{L \times 2K}$, reduced dimension $\mathbf{A}$ matrix (AWGN)	37
$\mathbf{B} \in \mathbb{C}^{2K \times 2K}$, matrix of received complex amplitudes	13
$\boldsymbol{\beta} \in \mathbb{R}^{2RM}$, vector of $\boldsymbol{\beta}(m)$ vectors	94
β_k , k th user's received complex amplitude (AWGN)	12
$\boldsymbol{\beta}_k(m) \in \mathbb{R}^{R_k}$, vector of $\beta_{k,r}(m)$ terms	94
$\beta_{k,r}(m)$, complex amplitude of k th user's r th ray (fading)	91
$\boldsymbol{\beta}(m) \in \mathbb{R}^R$, vector of $\boldsymbol{\beta}_k(m)$ vectors	94
$\mathbf{B}_\gamma \in \mathbb{C}^{2K \times 2K}$, "derivative" of $\mathbf{B}$ with respect to $\boldsymbol{\gamma}$	62

$\mathbf{b}_{k,1}(\tau), \mathbf{b}_{k,2}(\tau) \in \mathbb{R}^{QN}$, k th user's MUSIC test vectors (AWGN)	66
$\mathbf{b}_{k,1}(\tau), \mathbf{b}_{k,2}(\tau) \in \mathbb{R}^{QN}$, k th user's MUSIC test vectors (fading)	99
$b_k(t)$, one period of the k th user's code waveform	10
$\mathbf{B}' \in \mathbb{C}^{2(K-1) \times 2(K-1)}$, matrix of received complex amplitudes from the interfering users (AWGN)	24
$\mathbf{B}_\theta \in \mathbb{C}^{2K \times 2K}$, "derivative" of $\mathbf{B}$ with respect to θ	62
$\mathbf{c}_k \in \mathbb{R}^{QN}$, k th user's code vector	13
$\text{CRB}(\tau) \in \mathbb{R}^{R \times R}$, Cramér-Rao bound on estimators of τ (fading)	97
$\hat{d}_1(m) \in \{-1, 1\}$, hard decision of k th user's m th bit	21
δ_k , fractional part of τ_k ; $\delta_k = \tau_k - p_k T_i$	13
$\delta_{k,r}$, fractional part of $\tau_{k,r}$; $\delta_{k,r} = \tau_{k,r} - p_{k,r} T_i$	92
$d_k(m) \in \{-1, 1\}$, k th user's data sequence	10
$\mathbf{d}(m) \in \{-1, 1\}^{2K}$, data vector	22
$\mathbf{D}(r, s) \in \{0, 1\}^{QN \times QN}$, "permutation" matrix	13
$E_{b,k}$, k th user's energy per bit	12
$e(m)$, complex error signal	29
$\mathbf{E}_n \in \mathbb{R}^{QN \times (QN-2K)}$, matrix whose columns spans the noise subspace	15
$\mathbf{E}_s \in \mathbb{R}^{QN \times 2K}$, matrix whose columns spans the signal subspace	15
η_1 , desired user's asymptotic efficiency	26
$\bar{\eta}_1$, desired user's near-far resistance	26
$\mathbf{G} \in \{0, 1\}^{2K \times K}$	63
$\gamma \in \mathbb{R}^K$, vector of γ_k terms	13
$\hat{\gamma}^{\text{ML}} \in \mathbb{R}^K$, maximum likelihood estimate of γ (AWGN)	64
$\Gamma_k(m) \in \mathbb{C}^{2R_k \times R_k}$, matrix of $s_{k,r}(m)$ terms	96
$\gamma_k = \sqrt{P_k}$, magnitude of received complex amplitudes	14
$\mathbf{\Gamma}(m) \in \mathbb{C}^{2R \times R}$, matrix of $\Gamma_k(m)$ matrices	96

$g_k(n)$, impulse response to k th tap of gamma filter	43
J , mean square error	29
$\mathbf{J} \in \mathbb{R}^{(1+2RM+R) \times (1+2RM+R)}$, Fisher information matrix (fading)	95
$\mathbf{J} \in \mathbb{R}^{(1+3K) \times (1+3K)}$, Fisher information matrix (AWGN)	61
$\mathbf{J}_\infty \in \mathbb{R}^{(1+3K) \times (1+3K)}$, $\mathbf{J}$ as $M \rightarrow \infty$ (AWGN)	63
$J_{k,\text{MU}}(\tau)$, MUSIC cost function (fading)	99
$J_{\min}$, minimum mean-squared error	30
J_r , real mean-squared error	31
$\mathbf{J}_V \in \mathbb{R}^{(1+2RM+R) \times (1+2RM+R)}$, Fisher information matrix for the reduced dimension data (fading)	104
$\mathbf{J}_V \in \mathbb{R}^{(1+3K) \times (1+3K)}$, Fisher information matrix for the reduced dimension data (AWGN)	74
K , number of users	10
L , dimension of reduced dimension received vector	37
$\Lambda_s \in \mathbb{R}^{2K \times 2K}$, diagonal matrix of the largest eigenvalues of $\mathbf{R}$	16
λ_k , k th nonzero eigenvalue of $\mathbf{ASA}^*$	15
$\Lambda' \in \mathbb{R}^{2K \times 2K}$, diagonal matrix of the nonzero eigenvalues of $\mathbf{ASA}^*$	15
$\ell(\gamma, \theta, \tau)$, negative log-likelihood function	64
$\ln L(\mathbf{r})$, log-likelihood function (AWGN)	61
$\ln L(\mathbf{r})$, log-likelihood function (fading channel)	95
m_u , mean of $u_1(m)$ conditioned on $d_1(m)$	23
N , number of chips per bit	10
$N_0/2$, two-sided power spectral density of $n(t)$	11
$n(l)$, complex noise sequence	12
$\mathbf{n}(m) \in \mathbb{C}^{QN}$, noise vector, vector of $n(l)$ terms	12
$n(t)$, noise waveform	11

$\mathbf{n}_V(m) \in \mathbb{C}^L$, reduced dimension noise vector (AWGN)	37
$\mathbf{n}_V(m) = \mathbf{V}^* \mathbf{n}(m) \in \mathbb{C}^L$, reduced dimension noise vector (fading)	104
ω_c , carrier frequency	11
$\mathbf{P} \in \mathbb{R}^{2K \times 2K}$, diagonal matrix of P_k terms	63
$\mathbf{p} \in \mathbb{C}^{Q^N}$, crosscorrelation vector between $d_1(m)$ and $\mathbf{r}(m)$	29
$\mathbf{P}_{\mathbf{AZ}(m)} \in \mathbb{R}^{Q^N \times Q^N}$, projection matrix onto $\text{range}(\mathbf{AZ}(m))$	97
P_e , average error probability	23
$P_{e,GA}$, Gaussian approximation of P_e	23
$P_{e,V}$, average error probability for reduced dimension MMSE received	38
$\phi \in \mathbb{R}^{1+2RM+R}$, parameter vector (fading)	94
$\phi \in \mathbb{R}^{3K+1}$, parameter vector (AWGN)	60
$\hat{\phi} \in \mathbb{R}^{3K+1}$, estimate of ϕ	61
$\hat{\phi}^{\text{ML}} \in \mathbb{R}^{1+2MR+R}$, maximum likelihood estimate of ϕ (fading)	98
$\phi' \in \mathbb{R}^{2RM+R}$, parameter vector excluding σ^2 (fading)	103
P_k , k th user's received power; $P_1 = 1$	11
p_k , largest integer such that $p_k T_i \leq \tau_k$	13
$p_{k,r}$, largest integer such that $p_{k,r} T_i \leq \tau_k r$	92
$\Psi \in \mathbb{R}^{Q^N \times 2K}$, the “derivative” of $\mathbf{A}$ with respect to $\boldsymbol{\tau}$ (AWGN)	62
$\Psi \in \mathbb{R}^{Q^N \times 2R}$, “derivative” of $\mathbf{A}$ with respect to $\boldsymbol{\tau}$ (fading)	95
$\Psi_k \in \mathbb{R}^{Q^N \times 2R_k}$, “derivative” of $\mathbf{A}_k$ with respect to $\boldsymbol{\tau}_k$ (fading)	95
$\psi_{2k-1}, \psi_{2k} \in \mathbb{R}^{Q^N}$, derivative of $\mathbf{a}_{2k-1}$ and $\mathbf{a}_{2k}$ with respect to τ_k (AWGN)	62
$\psi_{k,2r}, \psi_{k,2r-1} \in \mathbb{R}^{Q^N}$, derivative of $\mathbf{a}_{2k-1}$ and $\mathbf{a}_{2k}$ with respect to $\tau_{k,r}$ (fading)	95
$\Psi_V = \mathbf{V}^* \Psi \in \mathbb{C}^{L \times 2R}$, reduced dimension Ψ matrix (fading)	104
$\mathbf{p}_V \in \mathbb{C}^L$, crosscorrelation vector between $d_1(m)$ and $\mathbf{r}_V(m)$	37
Q , oversampling factor	11

$Q(x)$, Q-function	23
R , the total number of rays from all users	92
$\mathbf{R} \in \mathbb{R}^{QN \times QN}$, correlation matrix for $\mathbf{r}(m)$	14
$\mathbf{r} \in \mathbb{C}^{MQN}$, vector of $\mathbf{r}(m)$ vectors	61
$r_I(l)$, real, or in-phase, part of $r(l)$	11
R_k , number of paths in k th user's channel	89
$\mathbf{r}_k(m) \in \mathbb{C}^{QN}$, k th user's contribution to $\mathbf{r}(m)$ (AWGN)	12
$\mathbf{r}_k(m) \in \mathbb{C}^{QN}$, contribution to $\mathbf{r}(m)$ from the k th user (fading)	91
$r(l) = r_I(l) + jr_Q(l)$, complex received sequence	11
$\mathbf{r}(m) \in \mathbb{C}^{QN}$, received vector, vector of $r(l)$ terms	12
$\hat{\mathbf{R}}_M \in \mathbb{C}^{QN \times QN}$, sample correlation matrix	46
$r_Q(l)$, imaginary, or quadrature-phase, part of $r(l)$	11
$r(t)$, received signal	11
$\mathbf{R}_V \in \mathbb{C}^{L \times L}$, correlation matrix for $\mathbf{r}_V(m)$ (AWGN)	37
$\mathbf{R}_V \in \mathbb{R}^{L \times L}$, correlation matrix for $\mathbf{r}_V(m)$ (AWGN)	16
$\mathbf{r}_V \in \mathbb{C}^{ML}$, vector of $\mathbf{r}_V(m)$ terms (fading)	104
$\mathbf{r}_V(m) = \mathbf{V}^* \mathbf{r}(m) \in \mathbb{C}^L$, reduced dimension received vector (fading)	103
$\mathbf{r}_V(m) = \mathbf{V}^* \mathbf{r}(m) \in \mathbb{C}^L$, reduced dimension received vector (AWGN)	16
$\mathbf{S} \in \mathbb{R}^{2K \times 2K}$, correlation matrix for $\mathbf{Bz}(m)$ (AWGN)	14
σ^2 , two-sided power spectral density of $n(l)$	12
σ_u^2 , variance of $u_1(m)$ conditioned on $d_1(m)$	23
$\mathbf{s}_k(m) \in \mathbb{C}^{2R_k}$, matrix $s_{k,r}(m)$ terms	91
$s_{k,2r-1}(m), s_{k,2r}(m)$, functions of $z_{2k-1}(m), z_{2k}(m)$, and $\beta_k(m)$ (fading)	91
$s_k(t)$, k th user's baseband signal	10
$\mathbf{s}(m) \in \mathbb{C}^{2R}$, vector of $\mathbf{s}_k(m)$ vectors (fading channel)	92
$\mathbf{s}(m) \in \mathbb{C}^{QN}$, AML auxiliary vector (AWGN)	65

T , symbol duration.....	10
$\mathcal{T}$, set of candidate MUSIC estimates (fading)	100
$\boldsymbol{\tau} \in \mathbb{R}^K$, vector of τ_k terms (AWGN)	13
$\boldsymbol{\tau} \in \mathbb{R}^R$, vector of $\boldsymbol{\tau}_k$ vectors (fading).....	95
$\boldsymbol{\tau}_k \in \mathbb{R}^{R_k}$, vector of $\tau_{k,r}$ terms (fading).....	95
$\hat{\boldsymbol{\tau}}^{\text{AML}} \in \mathbb{R}^R$, approximative maximum likelihood estimate of $\boldsymbol{\tau}$ (fading).....	99
$\hat{\boldsymbol{\tau}}^{\text{ML}} \in \mathbb{R}^K$, maximum likelihood estimate of $\boldsymbol{\tau}$ (AWGN).....	64
$\hat{\boldsymbol{\tau}}_k^{\text{MU}} \in \mathbb{R}^{R_k}$, MUSIC estimate of $\boldsymbol{\tau}_k$ (fading)	100
τ_k , k th user's propagation delay (AWGN).....	11
$\tau_{k,r}$, propagation delay for k th user's r th ray (fading).....	89
$T_c = T/N$, chip duration.....	10
$\boldsymbol{\theta} \in \mathbb{R}^K$, vector of θ_k terms (AWGN)	13
$\hat{\boldsymbol{\theta}}^{\text{ML}} \in \mathbb{R}^K$, maximum likelihood estimate of $\boldsymbol{\theta}$ (AWGN).....	64
θ_k , k th user's received phase (AWGN)	11
θ'_k , k th user's transmitted carrier phase	11
$\theta_{k,r}$, received phase of k th user's r th ray (fading).....	89
$T_i = T_c/Q$, integration time.....	11
$\mathbf{u} \in \mathbb{R}^{Q^N}$, the orthogonal projection of $\mathbf{c}_1$ onto $(\text{range}(\mathbf{A}'))^\perp$	28
$u_1(m) = \text{Re}\{y_1(m)\}$, soft decision of k th user's m th bit.....	22
$\mathbf{V} \in \mathbb{C}^{Q^N \times L}$, dimension reduction transform.....	16
$\mathbf{V}_{\text{CSFB}} \in \mathbb{R}^{Q^N \times L}$, CSFB dimension reduction transform	41
$\mathbf{V}_{\text{EGID}} \in \mathbb{C}^{Q^N \times L}$, EGID dimension reduction transform.....	46
$\mathbf{V}_\Gamma \in \mathbb{R}^{Q^N \times L}$, Gamma receiver dimension reduction transform.....	43
$\mathbf{V}_{\text{m}\Gamma} \in \mathbb{R}^{Q^N \times L}$, modified gamma receiver dimension reduction transform.....	44
$\mathbf{V}_o \in \mathbb{C}^{Q^N \times L}$, optimum dimension reduction transform	38
$\mathbf{V}_{\text{SDR}} \in \mathbb{R}^{Q^N \times L}$, SDR dimension reduction transform.....	44

$\mathbf{w} \in \mathbb{C}^{QN}$, receiver vector	21
$\mathbf{w}_d \in \mathbb{C}^{QN}$, receiver vector of the decorrelating receiver	28
$\mathbf{w}_o \in \mathbb{C}^{QN}$, receiver vector of the MMSE receiver	30
$\mathbf{w}_{o,V} \in \mathbb{C}^L$, receiver vector of reduced dimension MMSE receiver	37
$\mathbf{w}_s \in \mathbb{C}^{QN}$, receiver vector of the standard receiver	27
$y_1(m)$, receiver filter output	22
$\mathbf{z} \in \{-1, 0, 1\}^{2MK}$, vector of $\mathbf{z}(m)$ vectors	64
$\hat{\mathbf{z}}^{\text{ML}} \in \{-1, 0, 1\}^{2MK}$, maximum likelihood estimate of $\mathbf{z}$	64
$\mathbf{Z}_k(m) \in \{0, \pm 1\}^{2R_k \times R_k}$, matrix of $z_{2k-1}(m)$ and $z_{2k}(m)$ terms (fading)	96
$z_{2k-1}(m), z_{2k}(m) \in \{-1, 1\}$, functions of the k th user's data sequence	12
$\mathbf{Z}(m) \in \{-1, 0, 1\}^{2K \times K}$ matrix of $z_k(m)$ terms (AWGN)	62
$\mathbf{Z}(m) \in \{0, \pm 1\}^{2R \times R}$, matrix of $\mathbf{Z}_k(m)$ matrices (fading)	96
$\mathbf{z}(m) \in \{-1, 0, 1\}^{2K}$, vector of $z_{2k-1}(m)$ and $z_{2k}(m)$ terms	13
$\mathbf{z}'(m) \in \{-1, 0, 1\}^{2(K-1)}$, vector of $z_{2k-1}(m)$ and $z_{2k}(m)$ terms for $k = 2, \dots, K$ (AWGN)	24

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

DIRECT-SEQUENCE CODE-DIVISION MULTIPLE ACCESS SYSTEMS:
NEAR-FAR RESISTANT PARAMETER ESTIMATION AND DETECTION

By

Erik G. Ström

December 1994

Chairman: Fred J. Taylor
Cochairman: Scott L. Miller
Major Department: Electrical Engineering

This dissertation considers asynchronous direct-sequence code-division multiple-access communication systems operating over additive white Gaussian noise as well as fading channels. Particular attention is given to the near-far problem, which occurs when the received signal amplitudes from the concurrent users are very dissimilar.

A class of receivers, called linear single-bit single-user receivers, are defined, and their properties are analyzed. The *minimum mean-squared error* (MMSE) receiver is a particularly attractive receiver in this class. The MMSE receiver is investigated in great detail and is shown to be near-far resistant and optimum (in the error probability sense) when the interference is negligible. Optimum linear complexity reduction through dimension reduction is proposed and studied. Unfortunately, the optimum scheme requires side-information that is not readily available. Several practical, but suboptimum, schemes are therefore proposed and evaluated.

The parameter estimation problem—the problem of estimating each user's propagation delay, received amplitude, and phase—is introduced, and the *Cramér-Rao*

bound (CRB) is derived. The CRB, which bounds the accuracy with which the parameters may be estimated, serves as the optimality criterion. Optimum linear complexity reduction through dimension reduction is developed. Significant complexity reduction may be obtained; however, the optimum scheme again requires knowledge of unavailable side-information. The optimum *maximum-likelihood* (ML) estimator is presented, and several suboptimum schemes—with significantly lower complexity than the ML estimator—are proposed and studied. The proposed estimators are experimentally shown to be robust to the near-far problem, which, in turn, is shown to be fatal to the previously known methods.

The performance degradation of the decorrelating detector caused by propagation delay estimation errors is investigated in theory and quantified by simulations. It is shown that very accurate propagation delay estimation is needed to retain good performance.

CHAPTER 1

INTRODUCTION

1.1 Notation and Conventions

The *Kronecker delta function* is defined as

$$\delta_K(p) = \begin{cases} 1 & \text{if } p = 0 \\ 0 & \text{otherwise} \end{cases}$$

Matrices and vectors are typeset in bold face, for instance, $\mathbf{A}$ or $\mathbf{v}$. The (i, j) th element of a matrix $\mathbf{A}$ is denoted by $\mathbf{A}(i, j)$. With $\text{diag}(a_1, a_2, \dots, a_K)$ we mean the $K \times K$ diagonal matrix whose k th diagonal element is a_k , i.e.,

$$(\text{diag}(a_1, a_2, \dots, a_K))(k, k) = a_k$$

On the other hand, $\text{diag}(\mathbf{A})$ is a diagonal matrix with the same dimension and diagonal as $\mathbf{A}$,

$$(\text{diag}(\mathbf{A}))(i, j) = \mathbf{A}(i, i)\delta_K(i - j)$$

The transpose, conjugate transpose, and 2-norm of a vector $\mathbf{x}$ are denoted by $\mathbf{x}^T$, $\mathbf{x}^*$ and $\|\mathbf{x}\| = \sqrt{\mathbf{x}^* \mathbf{x}}$, respectively. We will frequently use conjugate transpose notation, $\mathbf{A}^*$, to indicate transpose of a real quantity $\mathbf{A}$. Conjugation of a scalar α may be denoted by either $\bar{\alpha}$ or α^* .

Two subspaces associated with a matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ are the *range* (*column space*)

$$\text{range}(\mathbf{A}) = \{\mathbf{A}\mathbf{x} : \mathbf{x} \in \mathbb{C}^n\}$$

and the *null space*

$$\text{null}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{C}^n : \mathbf{A}\mathbf{x} = \mathbf{0}\}$$

The *Fundamental Theorem of Linear Algebra* says that the *row space* of a matrix $\mathbf{A}$, i.e., $\text{range}(\mathbf{A}^*)$, is the orthogonal complement to the null space of $\mathbf{A}$,

$$\text{range}(\mathbf{A}^*) = (\text{null}(\mathbf{A}))^\perp$$

The correlation of two complex random variables x and y is defined as $E[\bar{x}y]$. This definition is in complete analogy with the real case [31]. However, some researchers in the communication field find it necessary to define the correlation as $E[\bar{x}y]/2$, when x and y are complex, and as $E[\bar{x}y]$, when x and y are real [36]. This author's opinion is that this convention introduces more confusion than convenience and should thus be avoided.

1.2 Introduction to Multiple-Access Communication Systems

A *multiple-access communication system* is a system that serves the communication needs of a population of concurrent users. An example of a multiple-access system is a cellular telephone system. We will only consider *digital communication* in this dissertation. As deduced from its name, a digital communication system transmits digital information, i.e., information that can be thought of as a stream of binary symbols (bits). A natural measure of the quality of the communication link is the average *probability of error*, or *bit error rate*, since the smaller the bit error rate, the less distortion is introduced in the transmitted data.

Radio-based multiple-access systems are typically required to operate inside a certain frequency band and below certain mean and peak powers, i.e., the transmitted signals must be bandlimited and restricted in peak and mean amplitude. The available bandwidth and allowed mean and peak powers can therefore be viewed as the *system resources*. On the other hand, we can quantify the users' communication demands by the needed *data rate*, or *symbol rate*, and the largest tolerable bit error rate. For example, a cellular telephone user might require that information is transmitted at a

data rate of 48×10^3 bits per second and with a probability of error less than 10^{-3} . A multiple-access system should be designed such that each user's communication demands are satisfied while using up a minimum of the system resources. An obvious performance measure for multiple-access systems is the *system capacity*, which is defined as the maximum number of users that can simultaneously communicate at a specified bit error rate.

It is common to divide multiple-access systems into two categories: *time-division multiple-access* (TDMA) systems and *frequency-division multiple-access* (FDMA) systems. In an FDMA system, the available bandwidth is divided into frequency channels, and each user is allowed to transmit continuously in its assigned channel. Conversely, in a TDMA system the time-axis is divided into *time slots*, and a user is allowed to use the entire system bandwidth during the assigned time slot.

Hybrids of FDMA and TDMA exist, e.g., the *American digital cellular* standard (ADC), in which the system bandwidth is divided into 30 kHz frequency channels, and each frequency channel is divided into three time slots. Thus, three TDMA users reside in each frequency channel.

A third form of multiple-access, *code-division multiple-access* (CDMA), has recently gained significant interest. In a CDMA system, all users are allowed to use the entire system bandwidth at all times. The separation between users is neither done in time nor in frequency, but rather with codes. The roots of these types of systems can be traced back to the 1920s [43]; however, until recently, CDMA has almost exclusively been used in military applications. CDMA systems can be divided into several subclasses, e.g., *frequency hopping* (FH-CDMA) systems and *direct-sequence* (DS-CDMA) systems. Note, however, that other types of CDMA and hybrids of FH-CDMA and DS-CDMA exist, e.g., the *joint tactical distribution system* (JTIDS) [43].

The manner in which FDMA, TDMA, and CDMA users are allocated system resources is illustrated in Figure 1.1. The total capacity of the system is represented by

a volume spanned by the three axis labeled time, frequency, and code. The shadowed region represents the fraction of the system resources that is allocated to a particular user.

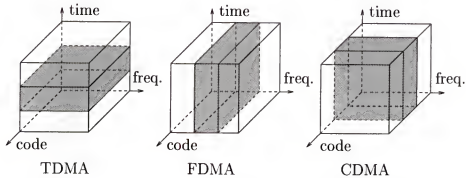


Figure 1.1: TDMA, FDMA and CDMA system resources allocation.

In FH-CDMA, the transmitted signal is a narrowband signal whose carrier frequency is changed (hopped) many times during the transmission. If the hopping rate is greater than the symbol rate, then this is called *fast* frequency hopping. Conversely, if the hopping rate is less than the symbol rate, then this is called *slow* frequency hopping. The hopping pattern is defined by a user-unique code, and the receiver (which is granted knowledge of the code) hops synchronously with the transmitter. An unauthorized receiver, however, will have difficulty tracking the hops (especially when fast hopping is used). The hopping makes the system more secure and robust against intentional or unintentional jamming. Every now and then, two or several users are going to hop simultaneously to the same frequency. This will garble the received signal and potentially destroy reception until the next hop. However, reliable communication can still be ensured if this situation does not occur too frequently and error correcting coding is used.

In DS-CDMA, the data signal—which is a rectangular waveform whose polarity is equal to the transmitted bit—is multiplied with a rectangular code waveform that

changes polarity at a rate much higher than the symbol rate. The multiplication of the data signal by the code waveform is called *spreading*. The standard receiver for DS-CDMA negates the spreading by multiplying the received signal with a copy of the desired user's code waveform. Consequently, this operation is called *despreading*. However, even after despreading, the signal will contain interference from the other concurrent users. The amount of interference is proportional to the crosscorrelations between the desired user's code waveform and the interfering users' code waveforms. Thus, reliable communication—when using the standard receiver—is possible only if the code waveforms have sufficiently low crosscorrelations.

In both FH-CDMA and DS-CDMA, the transmitted signal will occupy a bandwidth that is much larger than the bandwidth of the original data waveform. CDMA systems are therefore sometimes referred to as *spread spectrum* systems.

TDMA, FDMA, and CDMA have the same capacity if we model the channel as a synchronous *additive white Gaussian noise* (AWGN) channel [5]. However, this is a very simplistic and ideal model. For instance, a radio channel can seldom be modeled accurately as an AWGN channel. The propagation of a radio signal will often take several different paths to the receiver—so-called *multipath propagation*—which causes distortion to the received signal. A channel that distorts the transmitted signal is called a *fading* channel. A very general method to characterize a channel is to model the channel as a linear filter with an unknown time-varying impulse response [4]. The frequency response of the channel (i.e., the Fourier transform of the impulse response) typically has several nulls, or *fade dips*. We know, from elementary linear system theory, that the received signal is equal to the convolution of the transmitted signal and channel impulse response, and furthermore that the *power spectral density* (PSD) of the received signal is equal to the product of the transmitted signal PSD and the channel frequency response.

In FDMA systems, the transmitted signal—which is narrowband—can be com-

pletely wiped out by a fade dip, and communication is thereby made impossible. On the other hand, TDMA signals are more broadband than FDMA signals and are therefore not as vulnerable to fade dips. However, if the length of the impulse response is comparable to the symbol duration, the signal from a TDMA user will be noticeably smeared out in time, and parts of the signal will leak into the next user's time slot. The users will therefore interfere with each other. This problem can be alleviated by appending each time slot with a *guard interval*—a period of time during which no user is allowed to transmit. However, the use of guard intervals will decrease capacity, since no information can be exchanged during the guard interval.

Since the fading is time-variant, the location and depth of the fade dips will change in time. To combat fading, the system must use error correction coding or *diversity* techniques. With diversity we mean that several copies of a user's signal is transmitted over independently fading channels. We can then greatly reduce the bit error rate by combining the received signals from the channels. Diversity can be achieved in many ways. Firstly, we can transmit copies of the signal spaced in time over the same frequency channel (*time diversity*). If the spacing in time is greater than the time it takes for the channel impulse response to change significantly, we can assume that the signal copies fade independently. Secondly, we can use several antennas that are located sufficiently far away from each other to ensure independent propagation paths (*spatial diversity*). Thirdly, we can transmit the signal over several frequency channels that are spaced sufficiently in frequency (*frequency diversity*).

It can be argued that CDMA is more robust against fading than FDMA or TDMA [54, 16, 21, 36]. Heuristically, this can be done as follows. Since DS-CDMA signals are much more broadband than FDMA signals, fade dips cannot completely wipe out a DS-CDMA signal. Moreover, DS-CDMA is less sensitive to a long channel impulse response than TDMA, since the symbol rate is much slower in the DS-CDMA case. It can be seen that FH-CDMA has an inherent frequency diversity, since it hops

between frequency channels. A fade dip will therefore only affect the signal until the next hop. Furthermore, when a user has hopped from a frequency channel, the next time slot in that channel is not always used. Thus, the interference between users due to a long channel impulse response is less in an FH-CDMA system than in a TDMA system.

Another advantage of CDMA systems over TDMA and FDMA systems is the concept of *graceful degradation*. In FDMA and TDMA systems, the maximum number of users is fixed by the number of available frequency channels (FDMA) or time slots (TDMA). We simply cannot add a user when all channels or time slots are occupied. Furthermore, the bit error rate of a particular FDMA or TDMA user is independent of the number of concurrent users—at least in the ideal AWGN case. However, in CDMA there is no rigid bound on the maximum number of users. This is since the bit error rate is a function of the number of *active* users; by adding a user the performance is degraded for all active users, but the system will not collapse. Now, if the communication needs of the users are of bursty nature—for instance in a telephone situation when only one (well, ideally only one) user talks at a time—we can discontinue transmission during the time when communication is not needed. Capacity is thereby increased since there are fewer active users at any given time. Discontinuous transmission can be used to increase the capacity of FDMA and TDMA systems as well, but only with considerably more difficulty.

Yet another argument can be done for CDMA if we consider *cellular systems*. A cellular system is a system in which the geographical area to be covered is divided into cells, and the users in each cell communicate with a central *base station*. Cellular systems are motivated by two main reasons.

Firstly, the distance between a user and the base station is shorter in a multicell system than in a single cell system. The transmitted power from the mobile unit can thus be reduced. Reducing the transmitted power prolongs battery life and reduces

the physical size of the mobile device.

Secondly, and more importantly, is the concept of *frequency reuse*. By frequency reuse we mean that some cells in the system use the same frequency band. Capacity is thereby increased since the effective bandwidth of the system is increased. However, transmissions from frequency reusing cells will interfere with each other—so-called *cochannel interference*. Capacity is increased by shortening the distance between frequency reusing cells, however, cochannel interference is also increased. Thus, the smallest possible reuse distance, and, thereby, the largest capacity gain, is dependent on how robust the multiple-access system is toward interference. CDMA is less sensitive to cochannel interference than FDMA or TDMA, and CDMA might therefore be the best choice for a cellular system. Two cellular systems with frequency reuse are depicted in Figure 1.2. The shadowed cells are using the same frequencies and will therefore interfere with each other. Note that the frequency reuse distance is shorter in the left system than in the right system. Thus, if we can neglect the cochannel interference, the system to the left will have the larger capacity. However, the cochannel interference will also be larger in the leftmost system. Hence, the actual capacity difference is dependent on how sensitive the multiple-access system is to cochannel interference. As a matter of fact, the rightmost system might even have the larger capacity for some multiple-access methods.

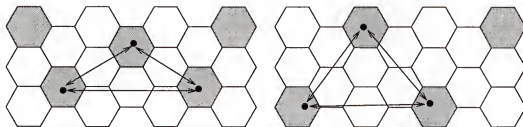


Figure 1.2: Cellular multiple-access systems with frequency reuse.

The question of which multiple-access system is the best is often subject to heated

debate. Each system has its own unique implementation problems, and the choice of multiple-access method is strongly dependent on application issues, such as the channel, the users' communication demands, and the applicable regulations, standards and laws.

Nice summaries of the early work on CDMA systems have been made by Pickholtz, Schilling and Milstein [33] and by Scholtz [42]. A thorough tutorial concerning the properties of code sequences has been written by Sarwate and Pursley [39]. Text books on CDMA include Dixon [10], Simon, Scholtz, Levitt and Omara [44], and Ziemer and Peterson [63]. Most modern books on digital communications contain a treatment on CDMA, e.g., Proakis [36], Lee and Messerschmitt [20], and Blahut [5].

The area of CDMA research is very active, and it is impractical to list all related work here. However, in the chapters to come, we will attempt to cite all papers that are relevant to the problem at hand.

1.3 Dissertation Outline

This dissertation is divided into two parts. The first part, Chapters 2–5, assumes a AWGN channel, and the second part, Chapters 6–7, assumes fading channels.

It may be argued that the first part is redundant since the AWGN channel is a special case of fading channels. However, it was decided to keep the treatment separate for two reasons. Firstly, the AWGN channel is much easier to analyze, and there are applications in which the channel is accurately modeled as an AWGN channel. Secondly, there is a much better understanding of the AWGN channel. For instance, there is not even a universally accepted model for a fading channel.

Parts of the original work presented in this dissertation have also been published at various IEEE conferences: [47, 52, 49, 51, 48, 50].

CHAPTER 2

SYSTEM MODEL FOR THE AWGN CHANNEL

The objective of this chapter is to formulate a convenient model for an asynchronous DS-CDMA communication system operating over an AWGN channel.

The original work in this chapter is threefold: (1) the formulation of the discrete-time complex vector model, as depicted in Figure 2.2 and defined by (2.17), (2) the derivation of the eigenvalue decompositions (2.29) and (2.37), and (3) the statement and proof of Theorem 2.3.

This chapter contains many fundamental definitions and algebraic manipulations, but very little discussion or new results. However, the author hopes that the dryness of this chapter will neither put the readers to sleep nor deter them from continuing to read the more interesting chapters to come.

2.1 Continuous-Time System Model

The system under consideration is modeled as an asynchronous K -user DS-CDMA system operating over an AWGN channel. The modulation scheme is binary phase-shift keying (BPSK) with bit duration T and chip duration $T_c = T/N$, where N is an integer. The code waveforms are assumed to be periodic with period T . As a general rule, a subscript k implies that the subscripted quantity is due to the k th user. For instance, a period of the k th user's code waveform is denoted by $b_k(t)$, where $b_k(t) = 0$ for $t \notin [0, T)$.

The baseband signal, $s_k(t)$, is formed by pulse amplitude modulating the data stream, $d_k(m) \in \{-1, 1\}$, with a period of the code waveform, i.e.,

$$s_k(t) = \sum_{m=-\infty}^{\infty} d_k(m)b_k(t - mT) \quad (2.1)$$

The transmitted signal is formed by multiplying $s_k(t)$ with the carrier $\sqrt{2P_k} \cos(\omega_c t + \theta'_k)$, where P_k is the power and θ'_k is the random carrier phase uniformly distributed in $[0, 2\pi)$. We assume, without loss of generality, that $P_1 = 1$ and $T_c = 1$.

The received signal may be written as

$$r(t) = \operatorname{Re} \left\{ \sum_{k=1}^K s_k(t - \tau_k) \sqrt{2P_k} \exp(j(\omega_c t + \theta_k)) \right\} + n(t) \quad (2.2)$$

where $\tau_k \in [0, T)$ is the unknown propagation delay and $\theta_k = \theta'_k - \omega_c \tau_k$. The noise waveform, $n(t)$, is a white Gaussian noise waveform with two-sided power spectral density $N_0/2$.

2.2 Discrete-Time Complex Vector System Model

The receiver front-end consists of a standard IQ-mixing stage followed by an integrate-and-dump section, as shown in Figure 2.1. The integration time, T_i , is defined as $T_i = T_c/Q$ where Q is an integer and is referred to as the *oversampling factor*.

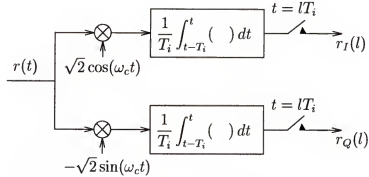


Figure 2.1: Receiver front-end.

Ignoring double frequency terms, the equivalent complex received sequence, $r(l) = r_I(l) + jr_Q(l)$, can be expressed as

$$r(l) = n(l) + \sum_{k=1}^K \sqrt{P_k} \exp(j\theta_k) \frac{1}{T_i} \int_{(l-1)T_i}^{lT_i} s_k(t - \tau_k) dt \quad (2.3)$$

where $n(l)$ is a zero-mean white complex Gaussian sequence with variance

$$\sigma^2 = \mathbb{E}[|n(l)|^2] = \frac{N_0}{T_i} = \frac{Q N N_0}{E_{b,1}} \quad (2.4)$$

where $E_{b,1} = T$ is the energy per bit for the first user.

Let the received vector during the m th bit interval, $\mathbf{r}(m) \in \mathbb{C}^{QN}$, and the noise vector, $\mathbf{n}(m) \in \mathbb{C}^{QN}$, be defined as

$$\mathbf{r}(m) = [r(mQN + QN) \quad r(mQN + QN - 1) \quad \cdots \quad r(mQN + 1)]^T \quad (2.5)$$

$$\mathbf{n}(m) = [n(mQN + QN) \quad n(mQN + QN - 1) \quad \cdots \quad n(mQN + 1)]^T \quad (2.6)$$

It is easy to show that the noise vector is a zero-mean complex Gaussian random vector with second moment matrices

$$\mathbb{E}[\mathbf{n}(p)\mathbf{n}^*(q)] = \sigma^2 \mathbf{I}_{QN} \delta_K(p - q), \quad \mathbb{E}[\mathbf{n}(p)\mathbf{n}^T(q)] = \mathbf{0} \quad (2.7)$$

After some straightforward calculations we can formulate the contribution from the k th user to $\mathbf{r}(m)$ as

$$\mathbf{r}_k(m) = [\mathbf{a}_{2k-1} \quad \mathbf{a}_{2k}] \begin{bmatrix} \beta_k & 0 \\ 0 & \beta_k \end{bmatrix} \begin{bmatrix} z_{2k-1}(m) \\ z_{2k}(m) \end{bmatrix} \quad (2.8)$$

where

$$z_{2k-1}(m) = \frac{1}{2}[d_k(m) + d_k(m-1)] \quad (2.9)$$

$$z_{2k}(m) = \frac{1}{2}[d_k(m) - d_k(m-1)] \quad (2.10)$$

$$\beta_k = \sqrt{P_k} \exp(j\theta_k) \quad (2.11)$$

From (2.8) we see that the k th user contributes $d_k(m)\beta_k\mathbf{a}_{2k-1}$, if $d_k(m) = d_k(m-1)$, or $d_k(m)\beta_k\mathbf{a}_{2k}$, if $d_k(m) = -d_k(m-1)$. The vectors $\{\mathbf{a}_{2k-1}, \mathbf{a}_{2k}\}$ are defined by the

k th user's propagation delay and code waveform as

$$\mathbf{a}_{2k-1} = \left[\frac{\delta_k}{T_i} \mathbf{D}(p_k + 1, 1) + \left(1 - \frac{\delta_k}{T_i} \right) \mathbf{D}(p_k, 1) \right] \mathbf{c}_k \quad (2.12)$$

$$\mathbf{a}_{2k} = \left[\frac{\delta_k}{T_i} \mathbf{D}(p_k + 1, -1) + \left(1 - \frac{\delta_k}{T_i} \right) \mathbf{D}(p_k, -1) \right] \mathbf{c}_k \quad (2.13)$$

where $\tau_k = p_k T_i + \delta_k$, such that p_k is an integer and $\delta_k \in [0, T_i]$, and $\mathbf{c}_k \in \mathbb{R}^{Q^N}$ is defined as

$$\mathbf{c}_k = [c_k(QN) \quad c_k(QN - 1) \quad \cdots \quad c_k(1)]^T \quad (2.14)$$

$$c_k(l) = \frac{1}{T_i} \int_{(l-1)T_i}^{lT_i} b_k(t) dt \quad (2.15)$$

The matrix $\mathbf{D}(r, s) \in \{0, 1\}^{Q^N \times Q^N}$ is defined in block form as

$$\mathbf{D}(r, s) = \begin{bmatrix} \mathbf{0} & \mathbf{I}_{Q^N - r} \\ s\mathbf{I}_r & \mathbf{0} \end{bmatrix} \quad (2.16)$$

We adopt the conventions that $\mathbf{D}(0, s) = \mathbf{I}_{Q^N}$ and $\mathbf{D}(Q^N, s) = s\mathbf{I}_{Q^N}$.

The expression for $\mathbf{r}(m)$ can be written in a more compact form as

$$\mathbf{r}(m) = \mathbf{n}(m) + \sum_{k=1}^K \mathbf{r}_k(m) = \mathbf{A}(\boldsymbol{\tau})\mathbf{B}(\boldsymbol{\gamma}, \boldsymbol{\theta})\mathbf{z}(m) + \mathbf{n}(m) \quad (2.17)$$

where

$$\mathbf{A}(\boldsymbol{\tau}) = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_{2K}] \in \mathbb{R}^{Q^N \times 2K} \quad (2.18)$$

$$\mathbf{B}(\boldsymbol{\gamma}, \boldsymbol{\theta}) = \text{diag}(\beta_1, \beta_1, \beta_2, \beta_2, \dots, \beta_K, \beta_K) \in \mathbb{C}^{2K \times 2K} \quad (2.19)$$

$$\mathbf{z}(m) = [z_1(m) \quad z_2(m) \quad \cdots \quad z_{2K}(m)]^T \in \{-1, 0, 1\}^{2K} \quad (2.20)$$

We observe that $\mathbf{A}(\boldsymbol{\tau})$ and $\mathbf{B}(\boldsymbol{\gamma}, \boldsymbol{\theta})$ are functions of the vectors $\boldsymbol{\tau}, \boldsymbol{\gamma}, \boldsymbol{\theta} \in \mathbb{R}^K$,

$$\boldsymbol{\tau} = [\tau_1 \quad \tau_2 \quad \cdots \quad \tau_K]^T \quad (2.21)$$

$$\boldsymbol{\theta} = [\theta_1 \quad \theta_2 \quad \cdots \quad \theta_K]^T \quad (2.22)$$

$$\boldsymbol{\gamma} = [\gamma_1 \quad \gamma_2 \quad \cdots \quad \gamma_K]^T \quad (2.23)$$

where $\gamma_k = \sqrt{P_k}$. The explicit dependence of $\mathbf{A}$ and $\mathbf{B}$ on the parameter vectors are dropped for notational convenience. We will assume that $\mathbf{A}$ and $\mathbf{B}$ have full rank. The $\mathbf{A}$ matrix will have full rank if and only if $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_K\}$ are linearly independent for all possible values of $\boldsymbol{\tau}$. This is obviously desirable for a DS-CDMA system since users otherwise can cancel each other's transmissions. The $\mathbf{B}$ matrix will be of full rank if $P_k > 0$ for all k , which clearly is the case. The final discrete-time complex vector model is depicted in Figure 2.2.

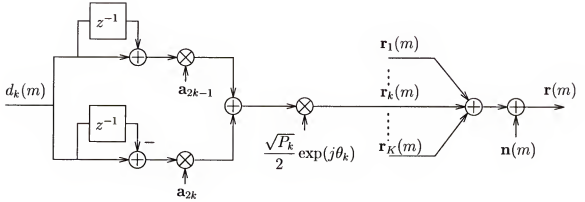


Figure 2.2: Discrete-time complex vector model

2.3 Properties of the Correlation Matrix

The correlation matrix for $\mathbf{r}(m)$ is

$$\mathbf{R} = \mathbf{E}[\mathbf{r}(m)\mathbf{r}^*(m)] = \mathbf{A}\mathbf{B}\mathbf{E}[\mathbf{z}(m)\mathbf{z}^*(m)]\mathbf{B}^*\mathbf{A}^* + \sigma^2\mathbf{I}_{QN} = \mathbf{A}\mathbf{S}\mathbf{A}^* + \sigma^2\mathbf{I}_{QN} \quad (2.24)$$

where τ_k , θ_k , and γ_k are considered to be unknown and deterministic parameters. We assume that the data streams consist of equally likely independent bits; furthermore, we assume that the noise is independent of the data streams. It is then easy to show that

$$\mathbf{S} = \mathbf{B}\mathbf{E}[\mathbf{z}(m)\mathbf{z}^*(m)]\mathbf{B}^* = \frac{1}{2} \text{diag}(P_1, P_1, \dots, P_K, P_K) \in \mathbb{R}^{2K \times 2K} \quad (2.25)$$

Note that $\mathbf{R}$ is symmetric and that $\mathbf{R}$ is positive definite if $\sigma^2 > 0$.

Since $\mathbf{A}\mathbf{S}\mathbf{A}^*$ is real, symmetric, positive semidefinite, and of rank $2K$, there exists an eigenvalue decomposition of $\mathbf{A}\mathbf{S}\mathbf{A}^*$ such that

$$\mathbf{A}\mathbf{S}\mathbf{A}^* = [\mathbf{E}_s \quad \mathbf{E}_n] \begin{bmatrix} \mathbf{\Lambda}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{E}_s \quad \mathbf{E}_n]^* \quad (2.26)$$

The matrices $\mathbf{E}_s \in \mathbb{R}^{QN \times 2K}$, $\mathbf{E}_n \in \mathbb{R}^{QN \times (QN - 2K)}$ are such that $[\mathbf{E}_s \quad \mathbf{E}_n] \in \mathbb{R}^{QN \times QN}$ is orthogonal. Furthermore, $\mathbf{\Lambda}' = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{2K}) \in \mathbb{R}^{2K \times 2K}$ is a diagonal matrix of the $2K$ nonzero and positive eigenvalues of $\mathbf{A}\mathbf{S}\mathbf{A}^*$.

We define the *signal subspace* to be the subspace spanned by the columns of $\mathbf{A}$, denoted by $\text{range}(\mathbf{A})$, and the *noise subspace* to be the orthogonal complement to the signal subspace. From (2.26), it follows that $\text{range}(\mathbf{E}_s) = \text{range}(\mathbf{A}\mathbf{S}\mathbf{A}^*)$, and the Theorem 2.1 below asserts that $\text{range}(\mathbf{A}) = \text{range}(\mathbf{A}\mathbf{S}\mathbf{A}^*)$. Thus, after proving the theorem, we conclude that the signal subspace is equal to $\text{range}(\mathbf{E}_s)$.

THEOREM 2.1 *Suppose $\mathbf{A} \in \mathbb{R}^{QN \times 2K}$ and $\mathbf{S} \in \mathbb{C}^{2K \times 2K}$, $QN \geq 2K$, are full rank matrices. Then $\text{range}(\mathbf{A}) = \text{range}(\mathbf{A}\mathbf{S}\mathbf{A}^*)$.*

PROOF: Since $(\text{range}(\mathbf{X}))^\perp = \text{null}(\mathbf{X}^*)$, it is sufficient to show that $\text{null}(\mathbf{A}^*) = \text{null}(\mathbf{A}\mathbf{S}\mathbf{A}^*)$, to prove the theorem.

Suppose $\mathbf{u} \in \text{null}(\mathbf{A}^*)$, then $\mathbf{A}^*\mathbf{u} = \mathbf{0}$ and

$$(\mathbf{A}\mathbf{S}\mathbf{A}^*)\mathbf{u} = \mathbf{0} \quad (2.27)$$

Therefore, $\mathbf{u} \in \text{null}(\mathbf{A}\mathbf{S}\mathbf{A}^*)$ or, equivalently, $\text{null}(\mathbf{A}^*) \subseteq \text{null}(\mathbf{A}\mathbf{S}\mathbf{A}^*)$.

Suppose $\mathbf{u} \in \text{null}(\mathbf{A}\mathbf{S}\mathbf{A}^*)$, then

$$\mathbf{0} = \mathbf{A}\mathbf{S}\mathbf{A}^*\mathbf{u} = (\mathbf{A}\mathbf{S})(\mathbf{A}^*\mathbf{u}) \quad (2.28)$$

Since $\mathbf{A}$ and $\mathbf{S}$ have full rank, it follows that $\mathbf{A}\mathbf{S}$ has full rank, and (2.28) then implies that $\mathbf{A}^*\mathbf{u} = \mathbf{0}$. Hence, $\mathbf{u} \in \text{null}(\mathbf{A}^*)$ or, equivalently, $\text{null}(\mathbf{A}\mathbf{S}\mathbf{A}^*) \subseteq \text{null}(\mathbf{A}^*)$, and the theorem follows. ■

It is easy to see that an eigenvector of $\mathbf{A}\mathbf{S}\mathbf{A}^*$ is also an eigenvector of $\mathbf{R}$. Thus, we can write an eigenvalue decomposition for $\mathbf{R}$ as

$$\begin{aligned}\mathbf{R} &= \mathbf{A}\mathbf{S}\mathbf{A}^* + \sigma^2 \mathbf{I}_{QN} = [\mathbf{E}_s \quad \mathbf{E}_n] \left(\begin{bmatrix} \Lambda' & 0 \\ 0 & 0 \end{bmatrix} + \sigma^2 \mathbf{I}_{QN} \right) [\mathbf{E}_s \quad \mathbf{E}_n]^* \\ &= [\mathbf{E}_s \quad \mathbf{E}_n] \begin{bmatrix} \Lambda_s & 0 \\ 0 & \Lambda_n \end{bmatrix} [\mathbf{E}_s \quad \mathbf{E}_n]^*\end{aligned}\quad (2.29)$$

where $\Lambda_s = \text{diag}(\lambda_1 + \sigma^2, \dots, \lambda_{2K} + \sigma^2)$ and $\Lambda_n = \sigma^2 \mathbf{I}_{QN-2K}$.

For reasons that will be explained later, we will now consider some of the properties of the correlation matrix for $\mathbf{r}_V(m) = \mathbf{V}^* \mathbf{r}(m)$. The matrix $\mathbf{V} \in \mathbb{C}^{QN \times L}$ is such that $\mathbf{V}^* \mathbf{V} = \mathbf{I}_L$, which implies that $L \leq QN$ and that $\mathbf{V}$ has full rank. Furthermore, we assume that the signal subspace is contained in the column space of $\mathbf{V}$, i.e., $\text{range}(\mathbf{E}_s) \subseteq \text{range}(\mathbf{V})$. The correlation matrix $\mathbf{R}_V \in \mathbb{R}^{L \times L}$ for $\mathbf{r}_V(m)$ is then

$$\mathbf{R}_V = E[\mathbf{r}_V(m) \mathbf{r}_V^*(m)] = \mathbf{V}^* \mathbf{A} \mathbf{S} \mathbf{A}^* \mathbf{V} + \sigma^2 \mathbf{V}^* \mathbf{V} = \mathbf{V}^* \mathbf{A} \mathbf{S} \mathbf{A}^* \mathbf{V} + \sigma^2 \mathbf{I}_L \quad (2.30)$$

The following theorems reveal some interesting properties of the correlation matrix.

THEOREM 2.2 *Suppose $\mathbf{V} \in \mathbb{C}^{QN \times L}$ is such that $\mathbf{V}^* \mathbf{V} = \mathbf{I}_L$ and $\text{range}(\mathbf{E}_s) \subseteq \text{range}(\mathbf{V})$ then*

$$\mathbf{V} \mathbf{V}^* \mathbf{u} = \mathbf{u} \quad (2.31)$$

for all $\mathbf{u} \in \text{range}(\mathbf{E}_s)$.

PROOF: Since $\mathbf{V}^* \mathbf{V} = \mathbf{I}_L$, the columns of $\mathbf{V}$ form an orthonormal basis for the subspace $\text{range}(\mathbf{V}) \supseteq \text{range}(\mathbf{E}_s)$. We can therefore express $\mathbf{u}$ as a linear combination of the columns of $\mathbf{V}$,

$$\mathbf{u} = \sum_{j=1}^L \alpha_j \mathbf{v}_j \quad (2.32)$$

where $\mathbf{v}_j$ is the j th column of $\mathbf{V}$. Hence,

$$\mathbf{V} \mathbf{V}^* \mathbf{u} = \mathbf{V} \mathbf{V}^* \sum_{j=1}^L \alpha_j \mathbf{v}_j = \sum_{j=1}^L \alpha_j \mathbf{V} \mathbf{V}^* \mathbf{v}_j = \sum_{j=1}^L \alpha_j \mathbf{V} \mathbf{e}_j = \sum_{j=1}^L \alpha_j \mathbf{v}_j = \mathbf{u} \quad (2.33)$$

where $\mathbf{e}_j$ is the j th unit vector. ■

THEOREM 2.3 Suppose an eigenvalue decomposition of $\mathbf{ASA}^*$ is given by (2.26), then an eigenvalue decomposition of $\mathbf{V}^*\mathbf{ASA}^*\mathbf{V}$, where $\mathbf{V} \in \mathbb{C}^{Q^{N \times L}}$ is such that $\mathbf{V}^*\mathbf{V} = \mathbf{I}_L$ and $\text{range}(\mathbf{E}_s) \subseteq \text{range}(\mathbf{V})$, can be written as

$$\mathbf{V}^*\mathbf{ASA}^*\mathbf{V} = [\mathbf{V}^*\mathbf{E}_s \quad \mathbf{E}_{n,V}] \begin{bmatrix} \Lambda' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} [\mathbf{V}^*\mathbf{E}_s \quad \mathbf{E}_{n,V}]^* \quad (2.34)$$

PROOF: In order to prove the theorem, it is sufficient to show that the columns of $\mathbf{V}^*\mathbf{E}_s$ are eigenvectors to $\mathbf{V}^*\mathbf{ASA}^*\mathbf{V}$ with corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_{2K}$.

Let $\mathbf{x}_j = \mathbf{V}^*\mathbf{u}_j$ where $\mathbf{u}_j$ is the j th column of $\mathbf{E}_s$. Since $\mathbf{u}_j \in \text{range}(\mathbf{E}_s) \subseteq \text{range}(\mathbf{V}) = (\text{null}(\mathbf{V}^*))^\perp$ it follows that $\mathbf{x}_j \neq \mathbf{0}$. Consider

$$\mathbf{V}^*\mathbf{ASA}^*\mathbf{V}\mathbf{x}_j = \mathbf{V}^*\mathbf{ASA}^*\mathbf{V}\mathbf{V}^*\mathbf{u}_j = \mathbf{V}^*\mathbf{ASA}^*\mathbf{u}_j \quad (2.35)$$

since, according to Theorem 2.2, $\mathbf{V}\mathbf{V}^*\mathbf{u}_j = \mathbf{u}_j$. Now, since $\mathbf{u}_j$ is an eigenvector of $\mathbf{ASA}^*$ with eigenvalue λ_j , it follows that

$$\mathbf{V}^*\mathbf{ASA}^*\mathbf{u}_j = \lambda_j \mathbf{V}^*\mathbf{u}_j = \lambda_j \mathbf{x}_j \quad (2.36)$$

Thus, $\mathbf{x}_j$ is an eigenvector of $\mathbf{V}^*\mathbf{ASA}^*\mathbf{V}$ with eigenvalue λ_j for $j = 1, 2, \dots, 2K$. Now, since $\text{rank}(\mathbf{ASA}^*) = 2K$, we know that $\text{rank}(\mathbf{V}^*\mathbf{ASA}^*\mathbf{V}) \leq 2K$ and therefore $\mathbf{V}^*\mathbf{ASA}^*\mathbf{V}$ will have at least $L - 2K$ zero eigenvalues. Moreover, since we just showed that there are $2K$ nonzero eigenvalues, $\mathbf{V}^*\mathbf{ASA}^*\mathbf{V}$ will have exactly $L - 2K$ zero eigenvalues. The theorem follows if we let $\mathbf{E}_{n,V}$ be an $L \times (L - 2K)$ matrix whose columns are eigenvectors of $\mathbf{V}^*\mathbf{ASA}^*\mathbf{V}$ corresponding to the zero eigenvalues. ■

We note that an eigenvector to $\mathbf{V}^*\mathbf{ASA}^*\mathbf{V}$ with eigenvalue λ_j is also an eigenvector to $\mathbf{R}_V$ with eigenvalue $\lambda_j + \sigma^2$. Hence, we can write the an eigenvalue decomposition of $\mathbf{R}_V$ as

$$\mathbf{R}_V = [\mathbf{V}^*\mathbf{E}_s \quad \mathbf{E}_{n,V}] \begin{bmatrix} \Lambda_s & \mathbf{0} \\ \mathbf{0} & \Lambda_{n,V} \end{bmatrix} [\mathbf{V}^*\mathbf{E}_s \quad \mathbf{E}_{n,V}]^* \quad (2.37)$$

where $\Lambda_s = \text{diag}(\lambda_1 + \sigma^2, \dots, \lambda_{2K} + \sigma^2)$ and $\Lambda_{n,V} = \sigma^2 \mathbf{I}_{L-2K}$.

CHAPTER 3 RECEIVERS FOR THE AWGN CHANNEL

3.1 Introduction to DS-CDMA Receivers

The DS-CDMA receiver problem has enjoyed a lot of attention during the last decade, and a tremendous number of papers have been published. The following introduction is not to be treated as a comprehensive review of all research in this field but merely as an indicator of major trends. The terms *receiver* and *detector* will be used interchangeably to denote a device that produces an estimate of the transmitted bit sequence by processing the received signal. As mentioned in the preceding chapter, the optimality criterion of a detector is the probability of error (or, equivalently, bit error rate), i.e., the probability that the estimated bit is not equal to the transmitted bit.

A summary of the state-of-the-art in the early 1980s has been written by Pickholtz, Shilling and Milstein [33]. In 1986, Verdú found the optimum receiver for DS-CDMA systems operating over AWGN channels [57]. Verdú showed that a remarkable performance gain may be achieved if the optimum receiver is used instead of the standard matched filter receiver.

The front-end of the optimum receiver is a bank of K filters. Each filter is matched to a particular user's code waveform and is sampled with respect to that user's propagation delay. The front-end output is a *sufficient statistic* for detecting the data sequences from all K users [57]. This means that the receiver front-end will not discard any information about the data sequences while processing the received signal [6]. The maximum-likelihood estimates of the transmitted data sequences are obtained

by processing the sufficient statistic with a Viterbi algorithm [58]. However, the optimum receiver suffers from two problems: (1) exponentially growing complexity in the number of users and (2) the need for perfect side-information, i.e., knowledge of the number of users and of each user's code waveform, propagation delay, received phase, and amplitude. The commonly accepted opinion is that the optimum receiver has little practical value, mainly because of its complexity. This opinion motivates research aimed at devising suboptimum detectors that have two properties: (1) less complexity than the optimum receiver and (2) better performance than the standard receiver.

Lupas, a student of Verdú's, has proposed a detector called the *decorrelating receiver* [23, 24, 22]. The performance of the decorrelating receiver is almost always superior to that of the standard receiver, and complexity grows linearly with the number of users. However, significant side-information is needed, e.g., knowledge of the number of users and of each user's code waveform, propagation delay, and received phase. The decorrelating receiver is a linear receiver in the sense that the decision is made as the sign of a linear function of the received signal. We will further discuss the Lupas decorrelating receiver in Chapter 5.

A nonlinear method, called the *multistage receiver*, has been proposed by Varanasi and Aazhang [56]. We can view the multistage receiver as an *interference cancellation* scheme. To understand this concept, consider the output of the k th user's matched filter. The matched filter output will contain a term from the desired user (i.e., the k th user) and an interference term due to the other $K - 1$ users. Clearly, if we knew this interference term, we could simply subtract it from the matched filter output and form the estimate of the k th user's data signal as the sign of the interference-free filter output. This strategy, as a matter of fact, has optimum performance. However, in practice, the interference is unknown and must therefore be estimated. Thus, an interference cancellation receiver is defined as a receiver that subtracts the

received signal with an estimate of the interference before making the final decision. In addition to the multistage receiver, many other interference cancellation schemes have been proposed, see for instance [11, 61] and the references therein.

DS-CDMA receivers can be classified in many categories. As mentioned above, the receiver is either linear or nonlinear. We can also divide DS-CDMA receivers into single-user and multiuser detectors. A multiuser receiver detects several users' data signals at the same time, whereas a single-user receiver is concerned with only one particular user's data signal. Finally, we distinguish between single-bit and sequence detectors. A sequence detector accumulates the received signal for several bit durations and makes a decision of several bits at once; however, a single-bit detector makes its decisions on a bit-by-bit basis. According to these definitions, the standard receiver is a linear single-user single-bit detector; the optimum and multistage receivers are nonlinear multiuser sequence detectors; and the decorrelator receiver is a linear multiuser sequence detector.

The scope of this chapter is limited to the class of *linear single-bit single-user* receivers. In Section 3.2, this class of receivers is defined in terms of a *generic receiver*, and an expression for the error probability is derived. The *near-far problem* is introduced, and two performance measures, called *asymptotic efficiency* and *near-far resistance*, are defined and formulas for computing these are derived. The *standard*, *decorrelating*, and *minimum-mean square error* (MMSE) receivers, which are special cases of the generic receiver, are introduced in Sections 3.3, 3.4, and 3.5, respectively. Even though the decorrelating receiver studied in this chapter is the single-user single-bit version of the Lupas decorrelating receiver, the receivers are quite different and should not be confused with each other.

The properties of the MMSE receiver are derived and discussed in Section 3.5.2. The problem of complexity reduction of the MMSE receiver through *dimension reduction* is presented in Section 3.6; the optimum scheme is derived and several sub-

optimum methods are proposed. In Section 3.6.5, we present simulation results that compare the proposed dimension reduction strategies.

A very simple modification can be made to the MMSE receiver that lowers the complexity and also increases the performance. Technically, this modification renders the receiver nonlinear, and the modified receiver therefore falls outside the scope of this chapter. However, since many of the results presented here will apply for the modified receiver as well, a discussion of the modified receiver is included and can be found in Section 3.7.

Throughout this chapter we will assume that the receiver is synchronized with the desired user (user number 1). We can then, without loss of generality, set $\tau_1 = 0$. This implies that $\mathbf{a}_1 = \mathbf{a}_2 = \mathbf{c}_1$ and that the received vector from the desired user is $\mathbf{r}_1(m) = d_1(m)\beta_1\mathbf{c}_1$. The results in this chapter may easily be extended to the case of imperfect synchronization. Furthermore, the synchronization problem is treated in Chapter 4.

The original work of this chapter is (1) the derivation of the effective SNR for the MMSE receiver (3.44), (2) the proofs of the MMSE receiver asymptotic behavior found in Theorems 3.1, 3.2 and 3.3, (3) the formulation of previously proposed complexity reduction schemes as special cases of linear dimension reduction, (4) the formulation and proof of the optimum dimension reduction transform in Theorem 3.5, and (5) the formulation of the gamma, modified gamma, SDR, and EGID dimension reduction schemes.

3.2 Generic Linear Receiver

In this chapter, we will limit the discussion to linear single-bit single-user receivers. A linear single-bit single-user receiver forms the *soft decision*—the input to the threshold device—as the real part of the inner product of the received vector and a *receiver vector* $\mathbf{w} \in \mathbb{C}^{Q_N}$, as shown in Figure 3.1. The hard decision, $\hat{d}_1(m)$, is

then formed as the sign of the soft decision. Receivers of this class does attempt to make neither a sequence nor a multiuser detection and will therefore be suboptimum in nature [22].

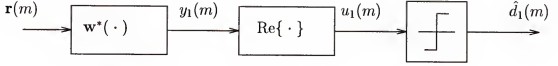


Figure 3.1: Generic receiver structure.

3.2.1 Probability of Error

The *filter output*, $y_1(m)$, is

$$y_1(m) = \mathbf{w}^* \mathbf{r}(m) = \mathbf{w}^* \mathbf{A} \mathbf{B} \mathbf{z}(m) + \mathbf{w}^* \mathbf{n}(m) \quad (3.1)$$

Now recall from (2.9), (2.10), and (2.20) that $\mathbf{z}(m)$ is a function of the data vector $\mathbf{d}(m) \in \{-1, 1\}^{2K}$ defined as

$$\mathbf{d}(m) = [d_1(m-1) \ \cdots \ d_K(m-1) \ d_1(m) \ \cdots \ d_K(m)]^T \quad (3.2)$$

Conditioned on $\mathbf{d}(m)$, or, equivalently, conditioned on $\mathbf{z}(m)$, the filter output will be a complex Gaussian random variable with mean

$$\mathbb{E}[y_1(m) \mid \mathbf{z}(m)] = \mathbf{w}^* \mathbf{A} \mathbf{B} \mathbf{z}(m) \quad (3.3)$$

and second moments

$$\mathbb{E}[|y_1(m) - \mathbf{w}^* \mathbf{A} \mathbf{B} \mathbf{z}(m)|^2 \mid \mathbf{z}(m)] = \mathbf{w}^* \mathbb{E}[\mathbf{n}(m) \mathbf{n}^*(m)] \mathbf{w} = \|\mathbf{w}\|^2 \sigma^2 \quad (3.4)$$

$$\mathbb{E}[(y_1(m) - \mathbf{w}^* \mathbf{A} \mathbf{B} \mathbf{z}(m))^2 \mid \mathbf{z}(m)] = \mathbf{w}^* \mathbb{E}[\mathbf{n}(m) \mathbf{n}^T(m)] \mathbf{w} = 0 \quad (3.5)$$

Thus, conditioned on $\mathbf{z}(m)$, the soft decision, $u_1(m) = \text{Re}\{y_1(m)\}$, will be a real Gaussian random variable with mean

$$\mathbb{E}[u_1(m) \mid \mathbf{z}(m)] = \text{Re}\{\mathbf{w}^* \mathbf{A} \mathbf{B} \mathbf{z}(m)\} \quad (3.6)$$

and variance

$$\mathbb{E}[(u_1(m) - \text{Re}\{\mathbf{w}^* \mathbf{A} \mathbf{B} \mathbf{z}(m)\})^2 \mid \mathbf{z}(m)] = \frac{\|\mathbf{w}\|^2 \sigma^2}{2} \quad (3.7)$$

We assume that the data sequences are statistically independent of each other and of the noise sequence. We also assume that all 2^{2K} possible realizations of $\mathbf{d}(m)$ are equally likely. This implies that we can write the average error probability as

$$\begin{aligned} P_e &= \Pr[\text{sgn}[u_1(m)] \neq d_1(m)] = \Pr[u_1(m) < 0 \mid d_1(m) = 1] \\ &= \sum_{\substack{\mathbf{d}(m) \in \{-1, 1\}^{2K} \\ d_1(m) = 1}} \Pr[u_1(m) < 0 \mid \mathbf{d}(m), d_1(m) = 1] \Pr[\mathbf{d}(m), d_1(m) = 1] \\ &= \frac{1}{2^{2K-1}} \sum_{\substack{\mathbf{d}(m) \in \{-1, 1\}^{2K} \\ d_1(m) = 1}} Q\left(\frac{\text{Re}\{\mathbf{w}^* \mathbf{A} \mathbf{B} \mathbf{z}(m)\}}{\sqrt{\|\mathbf{w}\|^2 \sigma^2 / 2}}\right) \end{aligned} \quad (3.8)$$

where the Q -function is defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-t^2/2) dt \quad (3.9)$$

The error probability can be estimated using the so-called *Gaussian approximation*. In essence, the Gaussian approximation means that the soft decision is modeled as a Gaussian random variable conditioned on $d_1(m)$. We denote the conditional mean and variance of the decision variable by m_u and σ_u^2 , respectively,

$$\begin{aligned} m_u &= \mathbb{E}[u_1(m) \mid d_1(m) = 1] = -\mathbb{E}[u_1(m) \mid d_1(m) = -1] \\ \sigma_u^2 &= \text{var}[u_1(m) \mid d_1(m) = 1] = \text{var}[u_1(m) \mid d_1(m) = -1] \end{aligned}$$

If $u_1(m)$ is Gaussian distributed conditioned on $d_1(m)$, it is straightforward to show that the error probability is

$$P_{e,\text{GA}} = Q\left(\frac{m_u}{\sqrt{\sigma_u^2}}\right) \quad (3.10)$$

We see that m_u^2/σ_u^2 plays the same role as the *signal-to-noise ratio* (SNR) in a single-user system, and we will therefore name this quantity the *effective SNR*. One should

keep in mind that the quality of the estimate (3.10) is dependent on how close the actual conditional distribution of $u_1(m)$ is to a Gaussian distribution.

To justify the Gaussian approximation, consider the soft decision. Recall, since $\tau_1 = 0$, that $\mathbf{a}_1 = \mathbf{a}_2 = \mathbf{c}_1$ and $z_1(m) + z_2(m) = d_1(m)$. The soft decision may be written as

$$\begin{aligned} u_1(m) &= \text{Re}\{y_1(m)\} \\ &= \underbrace{\text{Re}\{d_1(m) \exp(j\theta_1) \mathbf{w}^* \mathbf{c}_1\}}_{\text{signal part}} + \underbrace{\text{Re}\{\mathbf{w}^* \mathbf{A}' \mathbf{B}' \mathbf{z}'(m)\}}_{\text{multiple access interference}} + \underbrace{\text{Re}\{\mathbf{w}^* \mathbf{n}(m)\}}_{\text{noise part}} \end{aligned} \quad (3.11)$$

where

$$\mathbf{A}' = [\mathbf{a}_3 \quad \mathbf{a}_4 \cdots \mathbf{a}_{2K-1} \quad \mathbf{a}_{2K}] \in \mathbb{R}^{QN \times 2(K-1)} \quad (3.12)$$

$$\mathbf{B}' = \text{diag}(\beta_2, \beta_2, \dots, \beta_K, \beta_K) \in \mathbb{C}^{2(K-1) \times 2(K-1)} \quad (3.13)$$

$$\mathbf{z}'(m) = [z_3(m) \quad z_4(m) \quad \cdots \quad z_{2K-1}(m) \quad z_{2K}(m)]^T \in \{-1, 0, 1\}^{2(K-1)} \quad (3.14)$$

The first term in (3.11) is the desired signal, the second term is the *multiple access interference* (MAI), and the third term is an AWGN term. We see that the MAI is a sum of random variables

$$\text{Re}\{\mathbf{w}^* \mathbf{A}' \mathbf{B}' \mathbf{z}'(m)\} = \sum_{j=3}^{2K} (\text{Re}\{\mathbf{w}^* \mathbf{A}' \mathbf{B}'\})(j) z_j(m) \quad (3.15)$$

We are therefore tempted to invoke the *central limit theorem* that allows us to consider the MAI as a Gaussian random variable [31]. Furthermore, the MAI is independent of the noise and $d_1(m)$, since terms in the sum in (3.15) are independent of the noise and of $d_1(m)$.

Thus, if the central limit theorem is applicable (see Papoulis for a discussion [31]), we can model the MAI as Gaussian random variable. To compute the error probability, we need only find the variance and mean of the MAI. The mean of the MAI

is

$$\mathbb{E}[\text{Re}\{\mathbf{w}^* \mathbf{A}' \mathbf{B}' \mathbf{z}'(m)\}] = \text{Re}\{\mathbf{w}^* \mathbf{A}' \mathbf{B}'\} \mathbb{E}[\mathbf{z}'(m)] = 0 \quad (3.16)$$

and the conditional mean of $u_1(m)$ is then simply the signal part of (3.11),

$$m_u = \text{Re}\{d_1(m) \exp(j\theta_1) \mathbf{w}^* \mathbf{c}_1\} \quad (3.17)$$

The variance of the MAI is

$$\begin{aligned} \sigma_{\text{MAI}}^2 &= \mathbb{E}[\text{Re}^2\{\mathbf{w}^* \mathbf{A}' \mathbf{B}' \mathbf{z}'(m)\}] \\ &= \text{Re}\{\mathbf{w}^* \mathbf{A}' \mathbf{B}'\} \mathbb{E}[\mathbf{z}'(m) \mathbf{z}'^T(m)] \text{Re}^T\{\mathbf{w}^* \mathbf{A}' \mathbf{B}'\} \\ &= \frac{1}{2} \|\text{Re}\{\mathbf{w}^* \mathbf{A}' \mathbf{B}'\}\|^2 \end{aligned} \quad (3.18)$$

Since the MAI and noise are independent of each other and of $d_1(m)$, the conditional variance of $u_1(m)$ is the sum of the noise and MAI variances

$$\sigma_u^2 = \sigma_{\text{MAI}}^2 + \frac{\|\mathbf{w}\|^2 \sigma^2}{2} \quad (3.19)$$

and consequently we can approximate the error probability by

$$P_{e,\text{GA}} = Q\left(\frac{\text{Re}\{\exp(j\theta_1) \mathbf{w}^* \mathbf{c}_{V,1}\}}{\sqrt{\sigma_{\text{MAI}}^2 + \|\mathbf{w}\|^2 \sigma^2 / 2}}\right) \quad (3.20)$$

Note, however, that this equation is only valid if the central limit theorem is applicable.

3.2.2 The Near-Far Problem

Consider the MAI term found in (3.11) in more detail,

$$\mathbf{w}^* \mathbf{A}' \mathbf{B}' \mathbf{z}'(m) = \sum_{k=2}^K \sqrt{P_k} \exp(j\theta_k) [z_{2k-1}(m) \mathbf{w}^* \mathbf{a}_{2k-1} + z_{2k}(m) \mathbf{w}^* \mathbf{a}_{2k}] \quad (3.21)$$

The MAI is seen to be linear in $\sqrt{P_k}$. It is therefore impossible to detect $d_1(m)$ reliably if P_k becomes large and $\mathbf{w}$ is not orthogonal to $\mathbf{a}_{2k-1}$ and $\mathbf{a}_{2k}$. This is the

so-called *near-far problem* and is a serious threat to the performance of a DS-CDMA system.

The traditional solution to the near-far problem is to employ *power control*. The power control algorithm tries to keep all received powers at the same level by adjusting the transmitted powers. However, in practice this is a difficult task, and performance is severely degraded if the power control is imperfect [8].

The performance of a receiver is normally measured by the bit error rate. However, if we are mostly concerned with the near-far problem, a convenient performance measure is the *asymptotic efficiency* [22]. The asymptotic efficiency for the desired user is defined as

$$\eta_1 = \sup \left\{ 0 \leq r \leq 1 : \lim_{N_0/2 \rightarrow 0} \frac{P_e(N_0/2)}{Q\left(\sqrt{\frac{rE_{b,1}}{N_0/2}}\right)} < \infty \right\} \quad (3.22)$$

where $E_{b,k} = P_k T$ is the *energy per bit* for the k th user, and $P_e(N_0/2)$ is the probability of bit error given by (3.8). Recall that the error probability is a function of σ^2 , which in turn is a function of $N_0/2$; see (2.4). As seen from (3.22), the asymptotic efficiency is a number between 0 and 1. The Q-function in the denominator represents the error probability of a single-user system with an SNR of $r2E_{b,1}/N_0$. Thus, the asymptotic efficiency is a measure of how much extra bit energy is needed by the multiuser system to achieve the same bit error probability as a single-user system, in the limit as the noise vanishes. A nonzero asymptotic efficiency implies that $P_e(N_0/2)$ decays exponentially as $N_0/2 \rightarrow 0$, since $Q(x)$ decays exponentially for large x . On the other hand, for a linear receiver, a zero asymptotic efficiency implies an *error floor*, i.e., as the SNR goes to infinity, the receiver will still have a nonzero error probability [22].

The asymptotic efficiency is in general a function of the received amplitudes of the interference or, equivalently, a function of the interfering users' received bit energies, $E_{b,k}$ for $k = 2, 3, \dots, K$. The *near-far resistance*, $\bar{\eta}_1$, is defined as the worst case

asymptotic efficiency over all possible received energies¹

$$\bar{\eta}_1 = \inf_{\substack{E_{b,k} \\ k \neq 1}} \eta_1 \quad (3.23)$$

If the near-far resistance is zero, the receiver will then have an irreducible error floor as $N_0/2 \rightarrow 0$ in the worst case interference scenario. On the other hand, a nonzero near-far resistance guarantees that the receiver can reliably detect the desired user, regardless of the interference situation, if $2E_{b,1}/N_0$ is large enough.

In essence, the asymptotic efficiency and near-far resistance are measures on the robustness of the receiver to the near-far problem.

3.3 Standard Receiver

The standard receiver is simply a filter matched to the desired user's code waveform. In our setting this means that $\mathbf{w} = \mathbf{w}_s = \exp(j\theta_1)\mathbf{c}_1$. The soft decision for the standard receiver is

$$u_1(m) = \text{Re} \left\{ d_1(m) \|\mathbf{c}_1\|^2 + \sum_{k=2}^K \exp(-j\theta_1) \mathbf{c}_1^* \mathbf{r}_k(m) + \exp(-j\theta_1) \mathbf{c}_1^* \mathbf{n}(m) \right\} \quad (3.24)$$

The standard receiver is only optimum, i.e., has the lowest probability of bit error among all linear single-bit single-user receivers, if the MAI term in (3.24) is zero. The MAI is zero if $K = 1$ or if $\mathbf{c}_1$ is orthogonal to $\mathbf{r}_k(m)$ for $k = 2, 3, \dots, K$. We dismiss the $K = 1$ case and note that the second condition is equivalent to that $\mathbf{c}_1$ is orthogonal to $\{\mathbf{a}_3, \mathbf{a}_4, \dots, \mathbf{a}_{2K}\}$. The orthogonality condition cannot in general be satisfied in an asynchronous system, since the interfering users' $\mathbf{a}$ -vectors are functions of the propagation delays, see (2.12) and (2.13).

It is easy to show that the standard receiver has zero near-far resistance by letting $P_k \rightarrow \infty$ for $k = 2, 3, \dots, K$ [22]. The standard receiver is therefore only useful in a low MAI scenario.

¹Note that the bar over $\bar{\eta}_1$ does not indicate complex conjugation. This notation for near-far resistance is adopted for consistency with other work in the CDMA field.

3.4 Single-Bit, Single-User Decorrelating Receiver

When the interference becomes dominant (i.e., when the noise disappears, $\sigma \rightarrow 0$ or as $P_k \rightarrow \infty$ for $k = 2, 3, \dots, K$), the optimum receiver is a *decorrelating receiver*. The receiver vector for a decorrelating receiver, $\mathbf{w}_d$, is chosen such that it is orthogonal to the interfering $\mathbf{a}$ -vectors while still having a nonzero inner product with $\mathbf{c}_1$. The interference term is thereby eliminated and the soft decision becomes

$$u_1(m) = \text{Re}\{\mathbf{w}_d^* \mathbf{A} \mathbf{B} \mathbf{z}(m) + \mathbf{w}_d^* \mathbf{n}(m)\} = \text{Re}\{d_1(m) \mathbf{w}_d^* \mathbf{c}_1 \exp(j\theta_1) + \mathbf{w}_d^* \mathbf{n}(m)\} \quad (3.25)$$

From the requirements on $\mathbf{w}_d$, we conclude that the decorrelating receiver is guaranteed to exist if and only if $\mathbf{c}_1 \notin \text{span}\{\mathbf{a}_3, \mathbf{a}_4, \dots, \mathbf{a}_{2K}\} = \text{range}(\mathbf{A}')$.

The error probability follows from (3.8) as

$$P_{e,d} = Q\left(\frac{\text{Re}\{\mathbf{w}_d^* \mathbf{c}_1 \exp(j\theta_1)\}}{\sqrt{\|\mathbf{w}_d\|^2 \sigma^2 / 2}}\right) \quad (3.26)$$

Performance is therefore maximized if we choose $\mathbf{w}_d$ as the vector that maximizes the effective SNR,

$$\frac{2 \text{Re}^2\{\mathbf{w}_d^* \mathbf{c}_1 \exp(j\theta_1)\}}{\sigma^2 \|\mathbf{w}_d\|^2} \quad (3.27)$$

Thus, the formal definition of the decorrelating receiver vector is

$$\mathbf{w}_d = \exp(j\theta_1) \arg \max_{\mathbf{w} \in (\text{range}(\mathbf{A}'))^\perp} \frac{\mathbf{w}^* \mathbf{c}_1}{\|\mathbf{w}\|^2} \quad (3.28)$$

and it is easy to show that

$$\mathbf{w}_d = C \exp(j\theta_1) \mathbf{u} \quad (3.29)$$

where C is an arbitrary nonzero real constant and $\mathbf{u}$ is the orthogonal projection of $\mathbf{c}_1$ onto $(\text{range}(\mathbf{A}'))^\perp$.

Note that the decorrelating receiver as defined here is very different from the Lupas decorrelating receiver [22]. For instance, the Lupas decorrelating receiver is a multiuser sequence detector, i.e., it detects all K users' data sequences for all times in a single decision. This receiver is described in more detail in Chapter 5.

3.5 Minimum Mean-Squared Error Receiver

In the above we mentioned that the standard receiver is optimum if the interference is negligible, and the decorrelating receiver is optimum if the interference is dominant. Both of these situations are rather extreme in the sense that a practical system would probably be operating in a situation where neither noise nor interference may be ignored. The optimum choice of the receiver vector is therefore dependent on both the SNR and the interference level. An adaptive receiver that can alter its receiver vector depending on the current scenario is therefore desirable. The MMSE receiver has this capability, and we will show below that the MMSE receiver degenerates to the standard receiver and the decorrelating receiver as the MAI and noise disappear, respectively. It is hard to trace the exact origin of the MMSE receiver. Many researchers have mentioned the receiver; however, no one claims to be the inventor [27, 25, 1, 32, 61].

A block diagram of the MMSE receiver is shown in Figure 3.2. The adaptive algorithm adjusts the receiver vector such that the *mean square error* (MSE), $J = E[|e(m)|^2]$, is minimized. The error signal, $e(m)$, is formed as the difference between the filter output and the desired user's bit. The receiver will therefore require that a known data sequence, or *training sequence*, is transmitted before any actual data can be transmitted. Once the mean square error is at an acceptable level, data transmission begins and the error signal is formed as the difference between the filter output and the hard decision. These modes are referred to as *training* and *decision directed* modes, respectively.

It is well-known, see Haykin [15], that the MSE can be written as

$$J(\mathbf{w}) = E[|e(m)|^2] = E[d_1^2(m)] - \mathbf{p}^* \mathbf{w} - \mathbf{w}^* \mathbf{p} + \mathbf{p}^* \mathbf{R} \mathbf{p} \quad (3.30)$$

where $\mathbf{R}$ is the correlation matrix defined in (2.24), and $\mathbf{p}$ is the crosscorrelation

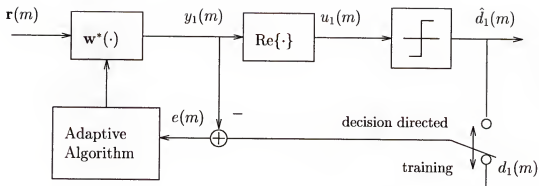


Figure 3.2: Block diagram of the minimum mean square error receiver.

vector

$$\mathbf{p} = E[d_1(m)\mathbf{r}(m)] = c_1 \exp(j\theta_1) \quad (3.31)$$

Differentiating (3.30) with respect to $\mathbf{w}$ and setting the result to zero yields the Wiener-Hopf equations

$$\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} = 2(\mathbf{R}\mathbf{w} - \mathbf{p}) = \mathbf{0} \quad (3.32)$$

If $\mathbf{R}$ is invertible, the minimizer of $J(\mathbf{w})$ and unique solution to (3.32) is the vector

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p} \quad (3.33)$$

and the resulting mean-squared error is

$$J_{\min} = J(\mathbf{w}_o) = 1 - \mathbf{p}^* \mathbf{R}^{-1} \mathbf{p} \quad (3.34)$$

3.5.1 Gaussian Approximation Error Probability for the MMSE Receiver

The general formula for the error probability (3.8) is, of course, valid for the MMSE receiver. However, we can also use the Gaussian approximation (3.20) to estimate the error probability, and, as indicated in Section 3.7.3, this approximation seems to be accurate. The effective SNR is related to the mean-squared error of the MMSE receiver as follows.

We define the *real mean-squared error* as

$$J_r = \mathbb{E}[\text{Re}^2\{e(m)\}] = \mathbb{E}[(d_1(m) - u_1(m))^2] \quad (3.35)$$

It is easy to show that $\mathbb{E}[\text{Re}^2\{e(m)\}] = \mathbb{E}[\text{Re}^2\{e(m)\} \mid d_1(m)]$. Hence, the real mean squared error is related to the effective SNR as

$$J_r = 1 - 2\mathbb{E}[u_1(m) \mid d_1(m)] + \mathbb{E}[u_1^2(m) \mid d_1(m)] = 1 - 2m_u + m_u^2 + \sigma_u^2 \quad (3.36)$$

and, thus, the effective SNR is

$$\frac{m_u^2}{\sigma_u^2} = \frac{m_u^2}{J_r - (1 - m_u)^2} \quad (3.37)$$

We now proceed to find expressions for J_r and m_u for the MMSE receiver.

$$J_r = \mathbb{E}[\text{Re}^2\{e(m)\}] = \frac{1}{4} \mathbb{E}[e^2(m) + \bar{e}^2(m) + 2|e(m)|^2] \quad (3.38)$$

We known from (3.34) that

$$\mathbb{E}[|e(m)|^2] = 1 - \mathbf{p}^* \mathbf{R}^{-1} \mathbf{p} \quad (3.39)$$

Consider $\mathbb{E}[e^2(m)]$,

$$\begin{aligned} \mathbb{E}[e^2(m)] &= \mathbb{E}[(d_1(m) - \mathbf{w}_o^* \mathbf{r}(m))^2] \\ &= 1 - 2\mathbb{E}[d_1(m) \mathbf{w}_o^* \mathbf{r}(m)] + \mathbf{w}_o^* \mathbb{E}[\mathbf{r}(m) \mathbf{r}^T(m)] \bar{\mathbf{w}}_o \\ &= 1 - 2\mathbf{w}_o^* \mathbf{p} + \frac{1}{2} \mathbf{w}_o^* \mathbf{A} \mathbf{B}^2 \mathbf{A}^T \bar{\mathbf{w}}_o \end{aligned} \quad (3.40)$$

It is easy to show that $\mathbb{E}[\bar{e}^2(m)] = \overline{\mathbb{E}[e^2(m)]}$ and

$$\mathbb{E}[\bar{e}^2(m) + e^2(m)] = 2\text{Re}\{\mathbb{E}[e^2(m)]\} = 2 - 4\mathbf{p}^* \mathbf{R}^{-1} \mathbf{p} + \text{Re}\{\mathbf{w}_o^* \mathbf{A} \mathbf{B}^2 \mathbf{A}^T \bar{\mathbf{w}}_o\} \quad (3.41)$$

We can now write J_r as

$$J_r = \frac{1}{4} [4 - 6\mathbf{p}^* \mathbf{R}^{-1} \mathbf{p} + \text{Re}\{\mathbf{w}_o^* \mathbf{A} \mathbf{B}^2 \mathbf{A}^T \bar{\mathbf{w}}_o\}] \quad (3.42)$$

Finally, we note that

$$\begin{aligned} m_u &= \mathbb{E}[\text{Re}\{\mathbf{w}_o^* \mathbf{r}(m)\} \mid d_1(m) = 1] \\ &= \text{Re}\{\mathbf{w}_o^* \mathbf{c}_1 \exp(j\theta_1)\} = \text{Re}\{\mathbf{w}_o^* \mathbf{p}\} = \mathbf{p}^* \mathbf{R}^{-1} \mathbf{p} \end{aligned} \quad (3.43)$$

By combining (3.37), (3.42), and (3.43), we can express the effective SNR for the MMSE receiver as

$$\frac{m_u^2}{\sigma_u^2} = \frac{4(\mathbf{p}^* \mathbf{R}^{-1} \mathbf{p})^2}{2\mathbf{p}^* \mathbf{R}^{-1} \mathbf{p} - 4(\mathbf{p}^* \mathbf{R}^{-1} \mathbf{p})^2 + \text{Re}\{\mathbf{w}_o^* \mathbf{A} \mathbf{B}^2 \mathbf{A}^T \bar{\mathbf{w}}_o\}} \quad (3.44)$$

3.5.2 Asymptotic Properties of the MMSE Receiver

We will now continue and prove two desirable properties of the MMSE receiver; the receiver degenerates to the matched filter, when the interference disappear, and to the decorrelating receiver, when noise disappear. In other words, the MMSE receiver will be the optimum receiver for both of these extreme situations.

The following theorem shows that if the interference is set to zero, the MMSE receiver degenerates to the matched filter, i.e., $\mathbf{w}_o = C \exp(j\theta_1) \mathbf{c}_1$, where C is an arbitrary nonzero real constant.

THEOREM 3.1 *Suppose that $\sigma > 0$. If $\mathbf{c}_1^* \mathbf{a}_{2k-1} = \mathbf{c}_1^* \mathbf{a}_{2k} = 0$ or $K = 1$, then the MMSE receiver vector is*

$$\mathbf{w}_o = \frac{\exp(j\theta_1)}{\sigma^2 + \|\mathbf{c}_1\|^2} \mathbf{c}_1 \quad (3.45)$$

PROOF: Recall that the correlation matrix can be written as

$$\mathbf{R} = \mathbf{A} \mathbf{S} \mathbf{A}^* + \sigma^2 \mathbf{I}_{QN} = \sigma^2 \mathbf{I}_{QN} + \mathbf{c}_1 \mathbf{c}_1^* + \sum_{k=2}^K \frac{P_k}{2} (\mathbf{a}_{2k-1} \mathbf{a}_{2k-1}^* + \mathbf{a}_{2k} \mathbf{a}_{2k}^*) \quad (3.46)$$

Consider

$$\mathbf{R}\mathbf{c}_1 = \sigma^2 \mathbf{c}_1 + \mathbf{c}_1 \mathbf{c}_1^* \mathbf{c}_1 + \sum_{k=2}^K \frac{P_k}{2} (\mathbf{a}_{2k-1} \mathbf{a}_{2k-1}^* + \mathbf{a}_{2k} \mathbf{a}_{2k}^*) \mathbf{c}_1 = (\sigma^2 + \|\mathbf{c}_1\|^2) \mathbf{c}_1 \quad (3.47)$$

Since $\mathbf{c}_1 \neq \mathbf{0}$, $\mathbf{c}_1$ is an eigenvector of $\mathbf{R}$ with eigenvalue $\lambda_1 = \sigma^2 + \|\mathbf{c}_1\|^2$. Thus,

$$\mathbf{R} \left(\frac{\exp(j\theta_1)}{\lambda_1} \mathbf{c}_1 \right) = \exp(j\theta_1) \mathbf{c}_1 = \mathbf{p} \quad (3.48)$$

Hence, $\exp(j\theta_1) \mathbf{c}_1 / \lambda_1$ satisfies the Wiener-Hopf equations (3.32), and since $\mathbf{R}$ is invertible, the MMSE receiver vector is unique and the theorem follows. ■

To show that the MMSE receiver degenerates to the decorrelating receiver as the noise goes to zero is a little bit more involved. The reason for this is that the correlation matrix is not invertible for $\sigma = 0$, and therefore the solution to the Wiener-Hopf equations (3.32) is not unique.

THEOREM 3.2 *Consider $\mathbf{R}$ and $J(\mathbf{w})$ as given by (2.24) and (3.30), respectively. Suppose $\mathbf{c}_1 \notin \text{span}\{\mathbf{a}_3, \mathbf{a}_4, \dots, \mathbf{a}_{2K}\}$. Then, as $\sigma \rightarrow 0$, the $\mathbf{w}_o \in \mathbb{C}^L$ that minimizes $J(\mathbf{w})$ is such that*

$$\lim_{\sigma \rightarrow 0} \mathbf{w}_o = \frac{\exp(j\theta_1)}{\|\mathbf{u}\|^2} \mathbf{u} \quad (3.49)$$

where $\mathbf{u}$ is the orthogonal projection of $\mathbf{c}_1$ onto $(\text{span}\{\mathbf{a}_3, \dots, \mathbf{a}_{2K}\})^\perp$

PROOF: The vector $\mathbf{w}_o$ satisfy the Wiener-Hopf equations $\mathbf{R}\mathbf{w} = \mathbf{p}$ where

$$\mathbf{R} = \mathbf{A}' \mathbf{S}' \mathbf{A}'^* + \mathbf{c}_1 \mathbf{c}_1^* + \sigma^2 \mathbf{I}_{QN} \quad (3.50)$$

and where $\mathbf{A}' = [\mathbf{a}_3 \ \dots \ \mathbf{a}_{2K}]$, $\mathbf{S}' = \frac{1}{2} \text{diag}(P_2, P_2, \dots, P_3, P_3)$, and $\mathbf{p} = \exp(j\theta_1) \mathbf{c}_1$. If $QN > 2K - 1$, then $\mathbf{R}$ is invertible if and only if $\sigma > 0$. For $\sigma = 0$ the Wiener-Hopf equations are

$$\mathbf{0} = \mathbf{R}\mathbf{w} - \mathbf{p} = \mathbf{A}' \mathbf{S}' \mathbf{A}'^* \mathbf{w} + \mathbf{c}_1 \mathbf{c}_1^* \mathbf{w} - \exp(j\theta_1) \mathbf{c}_1 \quad (3.51)$$

The columns of $\mathbf{A}'$ and $\mathbf{c}_1$ are linearly independent, and $\mathbf{S}'$ has full rank. Hence,

$$\mathbf{A}'^* \mathbf{w} = \mathbf{0} \Leftrightarrow \mathbf{w} \in \text{null}(\mathbf{A}'^*) = [\text{range}(\mathbf{A}')]^\perp \quad (3.52)$$

$$\mathbf{c}_1^* \mathbf{w} - \exp(j\theta_1) = 0 \quad (3.53)$$

Recall the eigenvalue decomposition of $\mathbf{R}$ (2.29),

$$\mathbf{R} = \mathbf{E}_s(\mathbf{A}' + \sigma^2 \mathbf{I}_{2K-1})\mathbf{E}_s^* + \sigma^2 \mathbf{E}_n \mathbf{E}_n^* \quad (3.54)$$

Since $\mathbf{c}_1 \in \text{range}(\mathbf{E}_s)$ it follows that $\mathbf{E}_n^* \mathbf{c}_1 = \mathbf{0}$, and when $\sigma > 0$, the solution to the Wiener-Hopf equations is

$$\mathbf{w}_o = \mathbf{R}^{-1} \mathbf{p} = \exp(j\theta_1) \mathbf{E}_s (\mathbf{A}' + \sigma^2 \mathbf{I}_{2K-1})^{-1} \mathbf{E}_s^* \mathbf{c}_1 \quad (3.55)$$

Clearly,

$$\lim_{\sigma \rightarrow 0} \mathbf{w}_o = \exp(j\theta_1) \mathbf{E}_s \mathbf{A}'^{-1} \mathbf{E}_s^* \mathbf{c}_1 \quad (3.56)$$

From this equation we observe that $\lim_{\sigma \rightarrow 0} \mathbf{w}_o \in \text{range}(\mathbf{E}_s)$, and since $\text{range}(\mathbf{E}_s) = \text{range}(\mathbf{A}) = \text{range}([\mathbf{c}_1 \quad \mathbf{A}'])$, it follows that

$$\text{range}(\mathbf{E}_s) = \text{range}(\mathbf{A}') + \text{range}(\mathbf{u}) \quad (3.57)$$

where $\mathbf{u}$ is the orthogonal projection of $\mathbf{c}_1$ onto $[\text{range}(\mathbf{A}')]^\perp$. Recall that, by assumption, $\mathbf{u} \neq \mathbf{0}$. Now, since $\lim_{\sigma \rightarrow 0} \mathbf{w}_o \in \text{range}(\mathbf{E}_s)$ and $\lim_{\sigma \rightarrow 0} \mathbf{w}_o \in [\text{range}(\mathbf{A}')]^\perp$, it is clear that $\lim_{\sigma \rightarrow 0} \mathbf{w}_o \in \text{range}(\mathbf{u})$. That is, for some complex scalar α ,

$$\lim_{\sigma \rightarrow 0} \mathbf{w}_o = \alpha \mathbf{u} \quad (3.58)$$

Consider (3.53),

$$\exp(j\theta_1) = \alpha \mathbf{c}_1^* \mathbf{u} = \alpha \|\mathbf{u}\|^2 \Rightarrow \alpha = \frac{\exp(j\theta_1)}{\|\mathbf{u}\|^2} \quad (3.59)$$

and the theorem follows. ■

We are now ready to derive an expression for the asymptotic efficiency for the MMSE receiver. As it turns out, the asymptotic efficiency is independent of the interfering users' received energies; therefore, for the MMSE receiver, asymptotic efficiency and near-far resistance are equivalent.

THEOREM 3.3 *The near-far resistance for a linear single-bit single-user receiver will be zero if $\mathbf{c}_1 \in \text{span}\{\mathbf{a}_3, \dots, \mathbf{a}_{2K}\}$. If $\mathbf{c}_1 \notin \text{span}\{\mathbf{a}_3, \dots, \mathbf{a}_{2K}\}$, then the asymptotic efficiency for the MMSE receiver will be*

$$\eta_1 = \frac{1}{QN\|\mathbf{w}_o\|^2} \quad (3.60)$$

where $\mathbf{w}_o$ is given by

$$\mathbf{w}_o = \frac{\exp(j\theta_1)}{\|\mathbf{u}\|^2} \mathbf{u} \quad (3.61)$$

and where $\mathbf{u}$ is the orthogonal projection of $\mathbf{c}_1$ onto $(\text{span}\{\mathbf{a}_3, \dots, \mathbf{a}_{2K}\})^\perp$

PROOF: The soft decision can be written as

$$\begin{aligned} y_1(m) &= d_1(m) \exp(j\theta_1) \mathbf{w}_o^* \mathbf{c}_1 + \mathbf{w}_o^* \mathbf{n}(m) \\ &\quad + \sum_{k=2}^K \sqrt{P_k} \exp(j\theta_k) (z_{2k-1}(m) \mathbf{w}_o^* \mathbf{a}_{2k-1} + z_{2k}(m) \mathbf{w}_o^* \mathbf{a}_{2k}) \end{aligned} \quad (3.62)$$

As the noise disappears and the interfering users' received powers goes to infinity, the soft decision will be dominated by the interference term unless $\mathbf{w}_o$ is orthogonal to $\{\mathbf{a}_3, \dots, \mathbf{a}_{2K}\}$. However, if $\mathbf{c}_1 \in \text{span}\{\mathbf{a}_3, \dots, \mathbf{a}_{2K}\}$, then $\mathbf{w}_o$ will be orthogonal to $\mathbf{c}_1$ as well, and the soft decision becomes zero. Clearly, this means that the near-far resistance of the receiver is zero. Conversely, if $\mathbf{c}_1 \notin \text{span}\{\mathbf{a}_3, \dots, \mathbf{a}_{2K}\}$, then the receiver may have a nonzero near-far resistance.

Suppose $\mathbf{c}_1 \notin \text{span}\{\mathbf{a}_3, \dots, \mathbf{a}_{2K}\}$, then Theorem 3.2 says that as $\sigma \rightarrow 0$,

$$\mathbf{w}_o = \frac{\exp(j\theta_1)}{\|\mathbf{u}\|^2} \mathbf{u} \quad (3.63)$$

where $\mathbf{u}$ the orthogonal projection of $\mathbf{c}_1$ onto $(\text{span}\{\mathbf{a}_3, \dots, \mathbf{a}_{2K}\})^\perp$. Hence,

$$y_1(m) = \frac{\exp(-j\theta_1)}{\|\mathbf{u}\|^2} \mathbf{u}^* \mathbf{c}_1 \exp(j\theta_1) d_1(m) = d_1(m) \quad (3.64)$$

Note that $y_1(m) = \mathbf{w}_o^* \mathbf{A} \mathbf{B} \mathbf{z}(m)$. Thus, the argument to the Q-functions in the expression for the error probability, (3.8), is seen to be independent of $\mathbf{d}(m)$. Thus, the

error probability can be written as

$$P_e(N_0/2) = Q\left(\sqrt{\frac{2}{\|\mathbf{w}_o\|^2\sigma^2}}\right) = Q\left(\sqrt{\frac{2E_{b,1}}{N_0} \frac{1}{QN\|\mathbf{w}_o\|^2}}\right) \quad (3.65)$$

since $\sigma^2 = QNN_0/E_{b,1}$. From the definition of asymptotic efficiency, (3.22), it follows that

$$\eta_1 = \frac{1}{QN\|\mathbf{w}_o\|^2} \quad (3.66)$$

Furthermore, since η_1 is independent of the interfering users' received energies, the asymptotic efficiency and the near-far resistance are equivalent. ■

3.6 Dimension Reduction

The MMSE receiver can be implemented as a QN -tap *finite impulse response* (FIR) filter, as shown in Figure 3.3. The FIR filter is fed the received sequence, (2.3), from the integrate-and-dump sections, and the filter is sampled at the symbol rate to form the filter output $y_1(m)$.

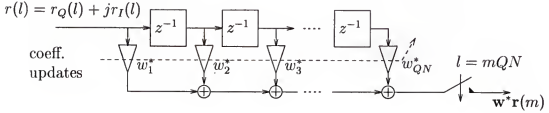


Figure 3.3: FIR implementation of the MMSE receiver.

It is desirable to reduce the length of the adaptive filter for several reasons. The computational burden on the adaptive algorithm will be reduced. The numerical stability is also increased, and the *excess mean square error*, i.e., the extra mean squared error added to $J_{\min}$ by the adaptation process, is decreased [15]. To reduce the filter length, we introduce a preprocessing stage in the form of a $QN \times L$ complex matrix $\mathbf{V}$, as shown in Figure 3.4. The received vector is premultiplied by $\mathbf{V}^*$, which

produces an L -vector. In other words, the dimensionality of the received vector is reduced from QN to L . We assume that $L \leq QN$ and that $\mathbf{V}$ has full rank.

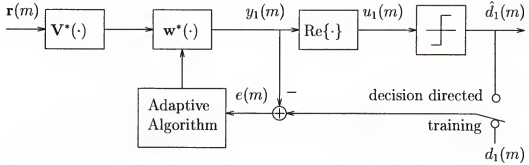


Figure 3.4: Reduced dimension MMSE receiver.

We denote the reduced dimensionality received vector with $\mathbf{r}_V(m) \in \mathbb{C}^L$;

$$\mathbf{r}_V(m) = \mathbf{V}^* \mathbf{r}(m) = \mathbf{V}^* \mathbf{A} \mathbf{B} \mathbf{z}(m) + \mathbf{V}^* \mathbf{n}(m) = \mathbf{A}_V \mathbf{B} \mathbf{z}(m) + \mathbf{n}_V(m) \quad (3.67)$$

where

$$\mathbf{A}_V = \mathbf{V}^* \mathbf{A} = [\mathbf{a}_{V,1} \quad \cdots \quad \mathbf{a}_{V,2K}] \in \mathbb{C}^{L \times 2K} \quad (3.68)$$

$$\mathbf{n}_V(m) = \mathbf{V}^* \mathbf{n}(m) \in \mathbb{C}^L \quad (3.69)$$

The correlation matrix for $\mathbf{r}_V(m)$ is

$$\mathbf{R}_V = \mathbb{E}[\mathbf{r}_V(m) \mathbf{r}_V^*(m)] = \mathbf{A}_V \mathbf{S} \mathbf{A}_V^* + \sigma^2 \mathbf{V}^* \mathbf{V} \quad (3.70)$$

The results obtained in Section 3.5 are still valid for the reduced dimension case if we make the appropriate substitutions and assumptions. The reduced dimension MMSE receiver vector, $\mathbf{w}_{o,V} \in \mathbb{C}^L$, is found as the solution to the Wiener-Hopf equations

$$\mathbf{R}_V \mathbf{w}_{o,V} = \mathbf{p}_V \quad (3.71)$$

where $\mathbf{p}_V \in \mathbb{C}^L$ is defined by

$$\mathbf{p}_V = \mathbb{E}[d_1(m) \mathbf{r}_V(m)] = \mathbf{V}^* \mathbf{p} = \mathbf{V}^* \mathbf{c}_1 \exp(j\theta_1) \quad (3.72)$$

The error probability is

$$P_{e,V} = \frac{1}{2^{2K-1}} \sum_{\substack{\mathbf{d}(m) \in \{-1,1\}^{2K} \\ d_1(m)=1}} Q \left(\frac{\operatorname{Re}\{\mathbf{w}_{o,V}^* \mathbf{A}_V \mathbf{B} \mathbf{z}(m)\}}{\sqrt{\|\mathbf{V} \mathbf{w}_{o,V}\|^2 \sigma^2 / 2}} \right) \quad (3.73)$$

We are naturally interested in finding the optimum dimension reduction transform. The optimum transform is defined as the transform that conserves the error probability, i.e., the transform $\mathbf{V}_o$ for which the MMSE receiver based on $\mathbf{V}_o^* \mathbf{r}(m)$ has the same error probability as the MMSE receiver based on $\mathbf{r}(m)$. The following lemma and theorem will give sufficient conditions for a transform to be optimum.

LEMMA 3.4 *Suppose that $\mathbf{E}_s$, $\mathbf{R}$, and $\mathbf{R}_V$ are defined by (2.29), (2.24), and (3.70), respectively, and that $\sigma > 0$. Furthermore, suppose that $\mathbf{V} \in \mathbb{C}^{L \times QN}$ is such that $\mathbf{V}^* \mathbf{V} = \mathbf{I}_L$, and $\operatorname{range}(\mathbf{E}_s) \subseteq \operatorname{range}(\mathbf{V})$. Then for $\mathbf{x}, \mathbf{y} \in \operatorname{range}(\mathbf{E}_s)$ and an integer p ,*

$$\mathbf{x}^* (\mathbf{R}^{-1})^p \mathbf{y} = (\mathbf{V}^* \mathbf{x})^* (\mathbf{R}_V^{-1})^p (\mathbf{V}^* \mathbf{y}) \quad (3.74)$$

PROOF: Note that $\mathbf{R}$ and $\mathbf{R}_V$ are invertible if $\sigma > 0$ and that $\mathbf{V}$ has full rank.

According to (2.37), we can write $(\mathbf{R}_V^{-1})^p$ as

$$(\mathbf{R}_V^{-1})^p = [\mathbf{V}^* \mathbf{E}_s \quad \mathbf{E}_{n,V}] \Lambda^{-p} [\mathbf{V}^* \mathbf{E}_s \quad \mathbf{E}_{n,V}]^* \quad (3.75)$$

where

$$\Lambda = \begin{bmatrix} \Lambda_s & \mathbf{0} \\ \mathbf{0} & \Lambda_{n,V} \end{bmatrix} \quad (3.76)$$

Let $\mathbf{u} \in \operatorname{range}(\mathbf{E}_s)$, then Theorem 2.2 allows us to write

$$(\mathbf{V}^* \mathbf{E}_s)^* (\mathbf{V}^* \mathbf{u}) = \mathbf{E}_s^* \mathbf{V} \mathbf{V}^* \mathbf{u} = \mathbf{E}_s^* \mathbf{u} \quad (3.77)$$

and since, according to Theorems 2.3 and 2.1, $\operatorname{range}(\mathbf{V}^* \mathbf{E}_s) = \operatorname{range}(\mathbf{V}^* \mathbf{A})$ and since $\mathbf{V}^* \mathbf{u} \in \operatorname{range}(\mathbf{V}^* \mathbf{E}_s)$,

$$\mathbf{E}_{n,V}^* (\mathbf{V}^* \mathbf{u}) = \mathbf{0} \quad (3.78)$$

Hence, for $\mathbf{x}, \mathbf{y} \in \text{range}(\mathbf{E}_s)$,

$$\begin{aligned}
(\mathbf{V}^* \mathbf{x})^* (\mathbf{R}_V^{-1})^p (\mathbf{V}^* \mathbf{y}) &= (\mathbf{V}^* \mathbf{x})^* [\mathbf{V}^* \mathbf{E}_s \quad \mathbf{E}_{n,V}] \Lambda^{-p} [\mathbf{V}^* \mathbf{E}_s \quad \mathbf{E}_{n,V}]^* (\mathbf{V}^* \mathbf{y}) \\
&= [(\mathbf{V}^* \mathbf{x})^* \mathbf{V}^* \mathbf{E}_s \quad (\mathbf{V}^* \mathbf{x})^* \mathbf{E}_{n,V}] \Lambda^{-p} \begin{bmatrix} (\mathbf{V}^* \mathbf{E}_s)^* (\mathbf{V}^* \mathbf{y}) \\ \mathbf{E}_{n,V}^* (\mathbf{V}^* \mathbf{y}) \end{bmatrix} \\
&= [(\mathbf{E}_s^* \mathbf{x})^* \quad \mathbf{0}] \Lambda^{-p} \begin{bmatrix} \mathbf{E}_s^* \mathbf{y} \\ \mathbf{0} \end{bmatrix} \\
&= \mathbf{x}^* \mathbf{E}_s \Lambda_s^{-p} \mathbf{E}_s^* \mathbf{y}
\end{aligned} \tag{3.79}$$

Since $\mathbf{E}_n^* \mathbf{x} = \mathbf{E}_n^* \mathbf{y} = \mathbf{0}$ and by virtue of (2.29), we can write (3.79) as

$$\begin{aligned}
(\mathbf{V}^* \mathbf{x})^* (\mathbf{R}_V^{-1})^p (\mathbf{V}^* \mathbf{y}) &= \mathbf{x}^* [\mathbf{E}_s \quad \mathbf{E}_n] \begin{bmatrix} \Lambda_s^{-p} & \mathbf{0} \\ \mathbf{0} & \Lambda_n^{-p} \end{bmatrix} [\mathbf{E}_s \quad \mathbf{E}_n]^* \mathbf{y} \\
&= \mathbf{x}^* (\mathbf{R}^{-1})^p \mathbf{y}
\end{aligned} \tag{3.80}$$

and the lemma follows. ■

Now we are ready to formulate and prove the form of the optimum dimension reduction transform.

THEOREM 3.5 *Suppose $\sigma^2 > 0$. If $\mathbf{V} \in \mathbb{C}^{QN \times L}$ is such that $\text{range}(\mathbf{E}_s) \subseteq \text{range}(\mathbf{V})$ and $\mathbf{V}^* \mathbf{V} = \mathbf{I}_L$, then the MMSE receiver with dimension reduction transform $\mathbf{V}$ will have the same error probability as the MMSE receiver without dimension reduction.*

PROOF: Recall that the probability of error for an MMSE receiver with dimension reduction transform $\mathbf{V}$ can be written as

$$P_{e,V} = \frac{1}{2^{2K-1}} \sum_{\substack{\mathbf{d}(m) \in \{-1,1\}^{2K} \\ d_1(m)=1}} Q \left(\frac{\text{Re}\{\mathbf{w}_{o,V}^* \mathbf{V}^* \mathbf{A} \mathbf{B} \mathbf{z}(m)\}}{\sqrt{\|\mathbf{V} \mathbf{w}_{o,V}\|^2 \sigma^2 / 2}} \right) \tag{3.81}$$

where $\mathbf{w}_{o,V} = \mathbf{R}_V^{-1} \mathbf{p}_V$, $\mathbf{R}_V = \mathbf{V}^* \mathbf{R} \mathbf{V}$, $\mathbf{p}_V = \mathbf{V}^* \mathbf{p}$, and $\mathbf{p} = \exp(j\theta_1) \mathbf{c}_1$. For the MMSE receiver without dimension reduction, the error probability is

$$P_e = \frac{1}{2^{2K-1}} \sum_{\substack{\mathbf{d}(m) \in \{-1,1\}^{2K} \\ d_1(m)=1}} Q \left(\frac{\text{Re}\{\mathbf{w}_o^* \mathbf{A} \mathbf{B} \mathbf{z}(m)\}}{\sqrt{\|\mathbf{w}_o\|^2 \sigma^2 / 2}} \right) \tag{3.82}$$

In order to prove the theorem, it is therefore sufficient to show that

$$\operatorname{Re}\{\mathbf{w}_{o,V}^* \mathbf{V}^* \mathbf{A} \mathbf{B} \mathbf{z}(m)\} = \operatorname{Re}\{\mathbf{w}_o^* \mathbf{A} \mathbf{B} \mathbf{z}(m)\} \quad (3.83)$$

$$\|\mathbf{V} \mathbf{w}_{o,V}\|^2 = \|\mathbf{w}_o\|^2 \quad (3.84)$$

Consider $\mathbf{w}_{o,V}^* \mathbf{V}^* \mathbf{A} \mathbf{B} \mathbf{z}(m)$ and recall that $\mathbf{R}_V^{-1}$ is symmetric;

$$\mathbf{w}_{o,V}^* \mathbf{V}^* \mathbf{A} \mathbf{B} \mathbf{z}(m) = \mathbf{p}_V^* (\mathbf{R}_V^{-1})^* \mathbf{V}^* \mathbf{A} \mathbf{B} \mathbf{z}(m) = (\mathbf{V}^* \mathbf{p})^* \mathbf{R}_V^{-1} (\mathbf{V}^* \mathbf{A} \mathbf{B} \mathbf{z}(m)) \quad (3.85)$$

Note that $\mathbf{p} \in \operatorname{range}(\mathbf{E}_s)$ and $\mathbf{A} \mathbf{B} \mathbf{z}(m) \in \operatorname{range}(\mathbf{E}_s)$. According to Lemma 3.4 this means that

$$(\mathbf{V}^* \mathbf{p})^* \mathbf{R}_V^{-1} (\mathbf{V}^* \mathbf{A} \mathbf{B} \mathbf{z}(m)) = \mathbf{p}^* \mathbf{R}^{-1} \mathbf{A} \mathbf{B} \mathbf{z}(m) = \mathbf{w}_o^* \mathbf{A} \mathbf{B} \mathbf{z}(m) \quad (3.86)$$

and it is seen that (3.83) is satisfied. Finally, since $\mathbf{V}^* \mathbf{V} = \mathbf{I}_L$,

$$\begin{aligned} \|\mathbf{V} \mathbf{w}_{o,V}\|^2 &= \|\mathbf{V} \mathbf{R}_V^{-1} \mathbf{p}_V\|^2 = \|\mathbf{V} \mathbf{R}_V^{-1} \mathbf{V}^* \mathbf{p}\|^2 \\ &= \mathbf{p}^* \mathbf{V} (\mathbf{R}_V^{-1})^* \mathbf{V}^* \mathbf{V} \mathbf{R}_V^{-1} \mathbf{V}^* \mathbf{p} = (\mathbf{V}^* \mathbf{p})^* \mathbf{R}_V^{-2} (\mathbf{V}^* \mathbf{p}) \end{aligned} \quad (3.87)$$

Now, since $\mathbf{p} \in \operatorname{range}(\mathbf{E}_s)$, we can use Lemma 3.4 again and

$$\|\mathbf{V} \mathbf{w}_{o,V}\|^2 = (\mathbf{V}^* \mathbf{p})^* \mathbf{R}_V^{-2} (\mathbf{V}^* \mathbf{p}) = \mathbf{p}^* \mathbf{R}^{-2} \mathbf{p} = \|\mathbf{w}_o\|^2 \quad (3.88)$$

Thus, (3.84) is satisfied, and the theorem follows. ■

From this theorem we conclude that the optimum dimension reduction transform is not unique and requires knowledge of the signal subspace $\operatorname{range}(\mathbf{E}_s) = \operatorname{range}(\mathbf{A})$. Hence, we cannot compute the optimum transform in practice, since $\mathbf{A}$ is a function the unknown propagation delays, see (2.12) and (2.13). Furthermore, the smallest L we can use is equal to the dimensionality of the signal subspace; in other words, $L \geq \operatorname{rank}(\mathbf{A})$. If we assume that $\tau_1 = 0$, then $L \geq 2K - 1$. However, if we know $\mathbf{w}_o$, i.e., the receiver vector for the MMSE receiver without dimension reduction, we may

use the transform $\mathbf{V} = \mathbf{w}_o$. Clearly, if $\mathbf{w}_{o,V} = C$, where C is an arbitrary real nonzero constant, then the resulting receiver is equivalent to the full length receiver. Thus, by definition, the dimension reduction transform $\mathbf{V} = \mathbf{w}_o$ is optimum. However, this transform violates the conditions of Theorem 3.5. With this in mind, it seems like the optimum transform according to Theorem 3.5 is of limited value. On the other hand, if we would like to implement a multiuser detector, the optimum transformation matrix will be optimum for all users in the system and furthermore $\mathbf{w}_o$ is of course unknown.

3.6.1 Cyclically Shifted Filter Bank

Madhow and Honig have proposed several reduced dimension MMSE receivers [25]. The authors were mainly concerned with the near-far problem, and the structures were therefore compared in terms of the asymptotic efficiency. The best receiver among the ones considered is the so-called *cyclically shifted filter bank* (CSFB). A block diagram for the CSFB receiver is shown in Figure 3.5

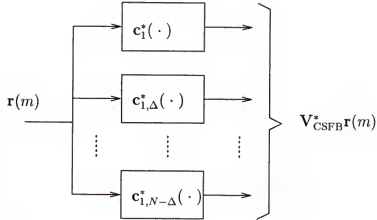


Figure 3.5: Block diagram of the cyclically shifted filter bank receiver.

The elements of $\mathbf{V}_{CSFB}^* \mathbf{r}(m)$ are found as the output of a bank of L filters, where each filter is matched to the desired vector cyclically shifted with $\Delta = QN/L$. We

assume, for convenience, that Δ is an integer. The CSFB dimension reduction transform, $\mathbf{V}_{\text{CSFB}} \in \mathbb{C}^{QN \times L}$, can be written as

$$\mathbf{V}_{\text{CSFB}} = [\mathbf{c}_1 \quad \mathbf{c}_{1,\Delta} \quad \cdots \quad \mathbf{c}_{1,QN-\Delta}] \quad (3.89)$$

where $\mathbf{c}_{1,m}$ denotes $\mathbf{c}_1$ cyclically shifted m steps. It is seen from (3.89) that $\mathbf{V}_{\text{CSFB}}$ cannot be optimum since all cyclical shifts of $\mathbf{c}_1$ are assumed to be linearly independent of $\{\mathbf{a}_3, \mathbf{a}_4, \dots, \mathbf{a}_{2K}\}$. Thus, $\mathbf{a}_j \notin \text{range}(\mathbf{V}_{\text{CSFB}})$ for any $j = 3, 4, \dots, 2K$, and the signal subspace will not be contained in $\text{range}(\mathbf{V}_{\text{CSFB}})$.

3.6.2 Gamma and Modified Gamma

The gamma filter MMSE receiver is depicted in Figure 3.6. The filter input is the received sequence $r(l) = r_I(l) + jr_Q(l)$, as defined by (2.3). The gamma filter is a *generalized feed-forward filter* found by replacing the delay elements in an ordinary tapped delay line with leaky integrators, or *gamma kernels* [35]. The gamma kernel is controlled by an adaptive parameter μ . The transfer function, $G(z)$, for a gamma kernel is

$$G(z) = \frac{\xi}{z - (1 - \mu)}, \quad \xi = \begin{cases} \mu & \text{for } 0 < \mu \leq 1 \\ 2 - \mu & \text{for } 1 < \mu < 2 \end{cases} \quad (3.90)$$

The kernel has a single pole at $1 - \mu$ and is therefore stable for $\mu \in (0, 2)$. It can be shown that μ and the weights $w_1, w_2, \dots, w_L$ can be adapted simultaneously such that the MSE is minimized [35].

The effect of the gamma filter is to perform dimension reduction and adaptation of the receiver vector simultaneously. We assume that the filter is cleared after each sampling of the filter. Thus, a filter cycle is as follows: the tap values are set to zero, QN samples of the received sequence are propagated through the cascaded gamma kernels, the resulting tap values are multiplied with the tap weights, and finally the filter is sampled.

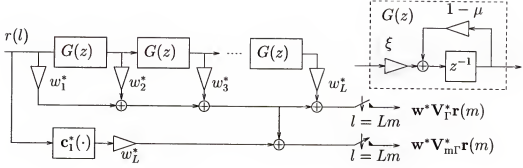


Figure 3.6: Block diagram of gamma and modified gamma receiver.

The dimension reduction transform is denoted by $\mathbf{V}_\Gamma \in \mathbb{R}^{Q^N \times L}$, and the (i, j) th element of $\mathbf{V}_\Gamma$ is

$$\mathbf{V}_\Gamma(i, j) = g_{j-1}(i) \quad (3.91)$$

where $g_k(n)$ is the impulse response from the input to the k th tap

$$g_k(n) = \begin{cases} \binom{n-1}{k-1} \xi^k (1-\mu)^{n-k} & \text{for } n \geq k \\ 0 & \text{for } n < k \end{cases} \quad (3.92)$$

As seen from (3.92), the dimension reduction is adaptive in nature. However, since $\mathbf{V}_\Gamma$ is controlled by only one parameter μ , the form of the dimension reduction transform is limited. For instance, it is easy to see that $\mathbf{V}_\Gamma$ cannot be optimum in the sense of Theorem 3.5. Nevertheless, we expect the gamma receiver to perform better than the CSFB when the system is interference dominated. This is since we have more freedom in the choice of dimension reduction transform in the gamma filter case. Conversely, we expect that the CSFB receiver will work better than the gamma receiver in a noise dominated scenario, since the gamma filter—as opposed to the cyclically shifted filter bank—can never be a matched filter. These heuristics were verified by the numerical results presented by Ström and Miller [47].

In an attempt to improve the performance of the gamma receiver, the structure was modified to include a tap that corresponds to the output of the matched filter.

The modified gamma receiver is depicted at the bottom of Figure 3.6, and the resulting dimension reduction transform, denoted by $\mathbf{V}_{\text{m}\Gamma} \in \mathbb{R}^{QN \times L}$, is

$$\mathbf{V}_{\text{m}\Gamma} = [\mathbf{V}_{\Gamma, L-1} \quad \mathbf{c}_1] \quad (3.93)$$

where $\mathbf{V}_{\Gamma, L-1}$ is the gamma transform with $L - 1$ taps.

The performance of the modified gamma receiver has been experimentally shown to be similar to the gamma receiver in a heavily loaded system and marginally better in a lighter loaded system [47].

3.6.3 Symmetric Dimension Reduction

The *symmetric dimension reduction* (SDR) scheme forms the elements of the reduced received vector as partial sums of the matched filter output. As seen from the block diagram in Figure 3.7, there are $\Delta = QN/L$ taps in each partial filter (Δ is again assumed to be an integer). If we let $w_1 = w_2 = \dots = w_L = \exp(j\theta_1)$, then the SDR receiver degenerates to the matched filter. Since the SDR has this property, it is expected that the symmetric dimension reduction receiver will perform similarly to the CSFB in a noise dominated scenario. We also note that the SDR receiver requires less computational complexity in the dimension reduction stage than the CSFB receiver.

The dimension reduction transform, $\mathbf{V}_{\text{SDR}} \in \mathbb{R}^{QN \times L}$, is found as

$$\mathbf{V}_{\text{SDR}} = [\mathbf{c}_{1,1:\Delta} \quad \mathbf{c}_{1,\Delta+1:2\Delta} \quad \dots \quad \mathbf{c}_{1,(L-1)\Delta+1:QN}] \quad (3.94)$$

where $\mathbf{c}_{1,r:p} \in \mathbb{R}^{QN}$, for $p \geq r$, denotes the vector

$$\mathbf{c}_{1,r:p} = [\underbrace{0 \quad \dots \quad 0}_{r-1 \text{ zeros}} \quad \mathbf{c}_1(r) \quad \mathbf{c}_1(r+1) \quad \dots \quad \mathbf{c}_1(p) \quad 0 \quad \dots \quad 0]^T \quad (3.95)$$

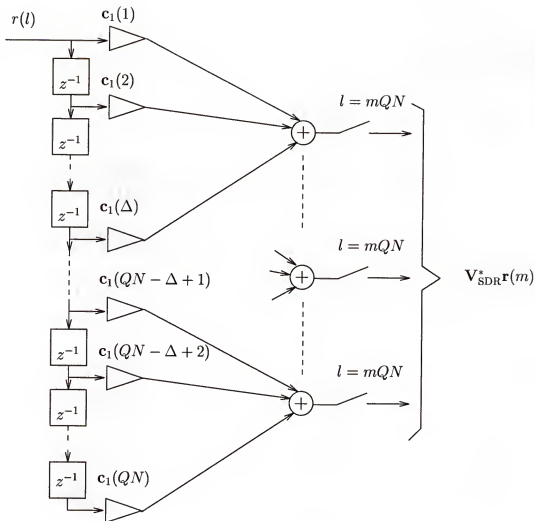


Figure 3.7: Block diagram of symmetric dimension reduction receiver.

3.6.4 Eigenvector Identification

The optimum transform cannot be computed directly in practice since the signal subspace, $\text{range}(\mathbf{E}_s)$, is unknown. However, we can certainly estimate $\text{range}(\mathbf{E}_s)$. Recall that $\mathbf{E}_s$ is a matrix whose columns are the eigenvectors corresponding to the $2K - 1$ largest eigenvalues of the correlation matrix $\mathbf{R}$. Now, since $\mathbf{R}$ is symmetric, $\mathbf{E}_s$ can be found as the $2K - 1$ first left singular vectors of $\mathbf{R}$. It is easy to see that $\mathbf{V} = \mathbf{E}_s$ satisfies the conditions for the optimum transform according to Theorem 3.5.

We can estimate $\mathbf{R} = \mathbb{E}[\mathbf{r}(m)\mathbf{r}^*(m)]$ with the *sample correlation matrix*

$$\hat{\mathbf{R}}_M = \frac{1}{M} \sum_{m=1}^M \mathbf{r}(m)\mathbf{r}^*(m) = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{W}}^* \quad (3.96)$$

where $\hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{W}}^*$ is the *singular value decomposition* (SVD) of $\hat{\mathbf{R}}_M$. $\mathbf{E}_s$ is now estimated by the $2K - 1$ first columns of $\hat{\mathbf{U}}$,

$$\mathbf{V}_{\text{EGID}} = [\hat{\mathbf{u}}_1 \quad \hat{\mathbf{u}}_2 \quad \cdots \quad \hat{\mathbf{u}}_{2K-1}] \quad (3.97)$$

This transform is called the *eigenvector identification* (EGID) dimension reduction transform.

The quality of the estimate increases with M ; as matter of fact, it can be shown that $\mathbf{V}_{\text{EGID}}$ is a *consistent* estimate of $\mathbf{E}_s$, that is, as $M \rightarrow \infty$, then $\mathbf{V}_{\text{EGID}} \rightarrow \mathbf{E}_s$. To compute $\mathbf{V}_{\text{EGID}}$, we can of course use an ordinary eigenvalue decomposition instead of an SVD.

3.6.5 Numerical Results

In this section we will present some results from Monte-Carlo simulations of different reduced dimension MMSE receivers. We will consider two performance measures: asymptotic efficiency and probability of bit error. As mentioned above, the asymptotic efficiency and the near-far resistance are equivalent for an MMSE receiver. The asymptotic efficiency is a number between zero and one. A zero near-far resistance implies that we will have an irreducible error as the noise power disappears. That is, we cannot reduce the error probability to an arbitrary level by increasing the signal-to-noise ratio. On the other hand, a nonzero near-far resistance implies that the error probability decreases exponentially with the signal-to-noise ratio when the noise power goes to zero. It is important to realize that a better near-far resistance does not imply a lower bit error probability in a noisy scenario. For instance, during heavy loading, that is, when $L \approx 2K$, the gamma and modified gamma receivers have

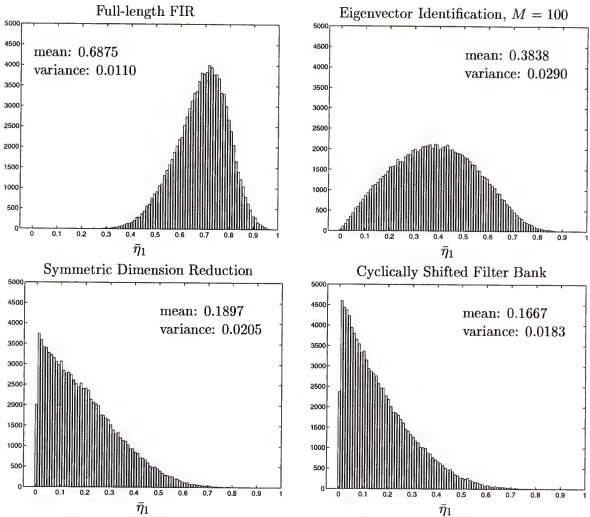


Figure 3.8: Histograms of the near-far resistance for 100,000 Monte-Carlo runs. The mean and variance indicated are computed as the sample mean and sample variance, respectively.

a larger near-far resistance than the cyclically shifted filter bank [47]. However, in a less loaded system or when we have a nonzero noise power, the cyclically shifted filter bank will outperform the gamma structures [47].

The simulated system was a $K = 5$ user system with $Q = 1$ and $N = 31$ chips per bit Gold sequences [39]. The dimension reduction was done by an $N \times L$ matrix where $L = 10$. This corresponds to rather heavily loaded system. In the bit error rate experiment, the signal-to-noise ratio, $2E_b/N_0$, was set to 10 dB for all users (i.e., perfect power control was assumed). To compute the EGID transform, according

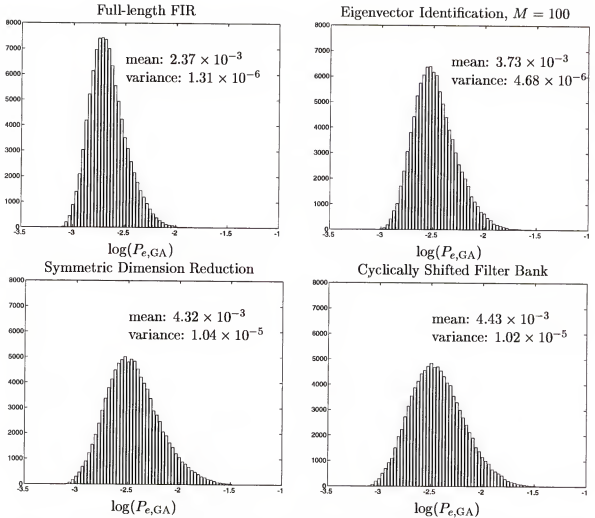


Figure 3.9: Histograms of the error probability for 100,000 Monte-Carlo runs. The mean and variance indicated are in linear units and are computed as the sample mean and sample variance, respectively.

to (3.97), we need to observe the received vector for M bit intervals. In the simulations $M = 100$ was used. Independent realizations of the propagation delays, data sequences, received phases, and noise vectors were drawn for each Monte-Carlo run. A total of 10^5 Monte-Carlo runs were done per histogram.

The plots presented in Figure 3.8 are histograms of the near-far resistance of the full-length receiver, cyclically shifted filter bank, eigenvector identification, and symmetric dimension reduction transforms. In Figure 3.9, histograms of the error probability for the same dimension reduction schemes are shown. The error proba-

bility was computed using the Gaussian approximation (3.20).

As seen from these plots, the full-length receiver is superior to the reduced dimension structures in both near-far resistance and error probability. The symmetric dimension reduction scheme is slightly better than the cyclically shifted filter bank. The EGID transform is clearly better than both the SDR and CSFB structures. For more extensive conclusions, see Section 3.8.

3.7 Linear Conjugate Linear MMSE Receiver

During the time this chapter was compiled, the author was made aware of a very simple modification to the MMSE receiver that achieves better performance with lower computational complexity [3]. The modification is the introduction of a nonlinearity, and thus the modified receiver does not belong to the class of linear receivers analyzed in this chapter. Nevertheless, this section was included since the modification is remarkably simple and it is believed that many of the results in presented previously in this chapter will hold for the modified receiver as well.

The underlying theory of this modification was developed by Brown and Crane as early as 1969 [7]. Aue and Reed named the modified receiver the *reduced linear-conjugate-linear Fourier series representation sequence estimator* (reduced LCL-FS-SE) [3]. For brevity and consistency with the notation in this dissertation, we choose to call the modified receiver the *linear conjugate linear MMSE* (LCL-MMSE) receiver.

3.7.1 Linear Conjugate Linear Filtering

If the correlation matrix $E[\mathbf{r}(m)\mathbf{r}^T(m)]$ is nonzero, then the introduction of a *conjugate branch* into the MMSE filter, as depicted in Figure 3.10, will increase the performance in the sense that the resulting MSE is lower than $J_{\min}$ as given by (3.34) [7]. Note that conjugation is a nonlinear operation.

We define the *complex correlation matrix* as $\mathbf{R}_c = E[\mathbf{r}(m)\mathbf{r}^T(m)]$. It easy to show,

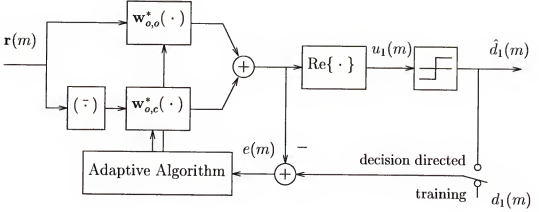


Figure 3.10: Block diagram of the linear conjugate linear MMSE receiver.

since $E[\mathbf{n}(m)\mathbf{n}^T(m)] = \mathbf{0}$ and $\mathbf{B}$ is diagonal, that

$$\mathbf{R}_c = E[\mathbf{r}(m)\mathbf{r}^T(m)] = \mathbf{A}\mathbf{B}E[\mathbf{z}(m)\mathbf{z}^T(m)]\mathbf{B}^T\mathbf{A}^T = \frac{1}{2}\mathbf{A}\mathbf{B}^2\mathbf{A}^T \quad (3.98)$$

$\mathbf{R}_c$ is clearly nonzero, and we therefore expect that the LCL-MMSE filter will perform better than the MMSE filter.

The error signal is formed as

$$e(m) = d_1(m) - \mathbf{w}_{o,o}^* \mathbf{r}(m) - \mathbf{w}_{o,c}^* \bar{\mathbf{r}}(m) = d_1(m) - [\mathbf{w}_{o,o}^* \quad \mathbf{w}_{o,c}^*] \begin{bmatrix} \mathbf{r}(m) \\ \bar{\mathbf{r}}(m) \end{bmatrix} \quad (3.99)$$

We can therefore view the LCL-MMSE filter as an augmented filter with filter vector $\mathbf{w}_a = [\mathbf{w}_{o,o}^T \quad \mathbf{w}_{o,c}^T]^T$ and filter input $\mathbf{r}_a(m) = [\mathbf{r}^T(m) \quad \bar{\mathbf{r}}^T(m)]^T$. The input can be written as

$$\mathbf{r}_a(m) = \mathbf{A}_a \mathbf{B}_a \mathbf{z}_a(m) + \mathbf{n}_a(m) \quad (3.100)$$

where

$$\begin{aligned} \mathbf{A}_a &= \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{bmatrix} \in \mathbb{R}^{2QN \times 4K} \\ \mathbf{B}_a &= \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{B}} \end{bmatrix} \in \mathbb{C}^{4K \times 4K} \\ \mathbf{z}_a(m) &= \begin{bmatrix} \mathbf{z}(m) \\ \mathbf{z}(m) \end{bmatrix} \in \{-1, 0, 1\}^{4K} \end{aligned}$$

$$\mathbf{n}_a(m) = \begin{bmatrix} \mathbf{n}(m) \\ \bar{\mathbf{n}}(m) \end{bmatrix} \in \mathbb{C}^{2Q_N}$$

It is straightforward to find the augmented filter vector as $\mathbf{w}_a = \mathbf{R}_a^{-1} \mathbf{p}_a$, where

$$\mathbf{R}_a = E[\mathbf{r}_a(m) \mathbf{r}_a^*(m)] = \begin{bmatrix} \mathbf{R} & \mathbf{R}_c \\ \bar{\mathbf{R}}_c & \mathbf{R} \end{bmatrix} \quad (3.101)$$

$$\mathbf{p}_a = E[d_1(m) \mathbf{r}_a(m)] = \begin{bmatrix} \mathbf{p} \\ \bar{\mathbf{p}} \end{bmatrix} \quad (3.102)$$

and where $\mathbf{R}$ and $\mathbf{p}$ is the correlation matrix and crosscorrelation vector defined by (2.24) and (3.31), respectively.

The complexity of the MMSE receiver, depicted in Figure 3.10, is roughly twice that of the ordinary MMSE, illustrated in Figure 3.2. Thus, whatever performance gain the LCL-MMSE receiver will give us might be offset by its increased complexity. However, we will show below that the LCL-MMSE receiver can be implemented with two *real* filters, rather than with two complex filters. The resulting complexity will actually be less than the ordinary MMSE receiver, since one complex multiplication usually requires more complexity than two real multiplies; therefore, one complex filter requires more complexity than two real filters.

Now consider the equation $\bar{\mathbf{p}}_a = \bar{\mathbf{R}}_a \mathbf{w}_a$

$$\bar{\mathbf{p}}_a = \begin{bmatrix} \bar{\mathbf{p}} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}} & \bar{\mathbf{R}}_c \\ \mathbf{R}_c & \bar{\mathbf{R}} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{w}}_{o,o} \\ \bar{\mathbf{w}}_{o,c} \end{bmatrix} \quad (3.103)$$

which can be rewritten as

$$\mathbf{p}_a = \begin{bmatrix} \mathbf{p} \\ \bar{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{R}} & \mathbf{R}_c \\ \bar{\mathbf{R}}_c & \bar{\mathbf{R}} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{w}}_{o,c} \\ \bar{\mathbf{w}}_{o,o} \end{bmatrix} = \mathbf{R}_a \begin{bmatrix} \bar{\mathbf{w}}_{o,c} \\ \bar{\mathbf{w}}_{o,o} \end{bmatrix} \quad (3.104)$$

since $\bar{\mathbf{R}} = \mathbf{R}$. Hence, since $\mathbf{R}_a \mathbf{w}_a = \mathbf{p}_a$,

$$\mathbf{w}_a = \begin{bmatrix} \mathbf{w}_{o,o} \\ \mathbf{w}_{o,c} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{w}}_{o,c} \\ \bar{\mathbf{w}}_{o,o} \end{bmatrix} \quad (3.105)$$

or $\mathbf{w}_{o,o} = \bar{\mathbf{w}}_{o,c}$. The error signal may therefore be written as

$$\begin{aligned} e(m) &= d_1(m) - \mathbf{w}_{o,o}^* \mathbf{r}(m) - \bar{\mathbf{w}}_{o,o}^* \bar{\mathbf{r}}(m) \\ &= d_1(m) - 2 \operatorname{Re}\{\mathbf{w}_{o,o}^* \mathbf{r}(m)\} \\ &= d_1(m) - 2 \left[\operatorname{Re}^T\{\mathbf{w}_{o,o}\} \operatorname{Re}\{\mathbf{r}(m)\} - \operatorname{Im}^T\{\mathbf{w}_{o,o}\} \operatorname{Im}\{\mathbf{r}(m)\} \right] \end{aligned} \quad (3.106)$$

In other words, the LCL-MMSE receiver can be implemented as shown in Figure 3.11, where $\mathbf{w}_{o,r} = 2 \operatorname{Re}\{\mathbf{w}_{o,o}\}$ and $\mathbf{w}_{o,i} = -2 \operatorname{Im}\{\mathbf{w}_{o,o}\}$. Note that this implementation only uses real filters, and consequently the computational complexity is reduced compared with the ordinary MMSE receiver (which requires complex filters).

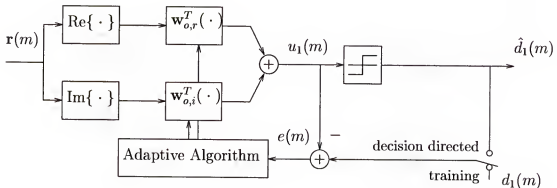


Figure 3.11: LCL-MMSE receiver using only real filters.

3.7.2 MMSE Receiver with A Priori Phase Knowledge

Suppose we know the desired user's phase, θ_1 . A natural modification to the MMSE receiver is then to remove the desired user's phase from the received vectors and process only the real part of the resulting vectors, as depicted in Figure 3.12. This receiver is called the *real MMSE* receiver.

The input to the filter is

$$\mathbf{r}_r(m) = \operatorname{Re}\{\mathbf{r}(m) \exp(-j\theta_1)\} = \mathbf{A}\mathbf{B}_r \mathbf{z}(m) + \mathbf{n}_r(m) \quad (3.107)$$

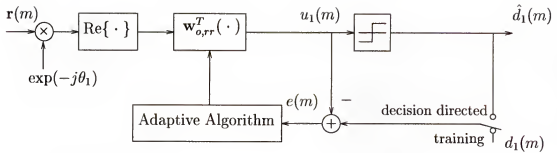


Figure 3.12: MMSE receiver using only real filtering.

where $\mathbf{n}_r(m) = \text{Re}\{\mathbf{n}(m) \exp(-j\theta_1)\}$ is a real Gaussian zero-mean random vector with correlation matrix

$$\mathbb{E}[\mathbf{n}_r(m) \mathbf{n}_r^T(m)] = \frac{\sigma^2}{2} \mathbf{I}_{QN} \quad (3.108)$$

and $\mathbf{B}_r \in \mathbb{R}^{2K \times 2K}$ is

$$\mathbf{B}_r = \text{Re}\{\mathbf{B} \exp(-j\theta_1)\} \quad (3.109)$$

The correlation matrix is

$$\mathbf{R}_r = \mathbb{E}[\mathbf{r}_r(m) \mathbf{r}_r^T(m)] = \frac{1}{2} [\mathbf{A} \mathbf{B}_r^2 \mathbf{A}^T + \sigma^2 \mathbf{I}_{QN}] \quad (3.110)$$

and the crosscorrelation vector is

$$\mathbf{p}_r = \mathbb{E}[d_1(m) \mathbf{r}_r(m)] = \mathbf{c}_1 \quad (3.111)$$

It is well-known that the minimum MSE is achieved by using the receiver vector $\mathbf{w}_{o,rr} = \mathbf{R}_r^{-1} \mathbf{p}_r$, and that the minimum MSE is

$$J_r = 1 - \mathbf{p}_r^T \mathbf{R}_r^{-1} \mathbf{p}_r \quad (3.112)$$

The decision variable is simply the output of the filter, and thus the mean is

$$m_u = \mathbb{E}[\mathbf{w}_{o,rr}^T \mathbf{r}_r(m) \mid d_1(m) = 1] = \mathbf{w}_{o,rr}^T \mathbf{c}_1 = \mathbf{w}_{o,rr}^T \mathbf{p}_r \quad (3.113)$$

The effective SNR for the real receiver can now be found as

$$\frac{m_u^2}{\sigma_u^2} = \frac{m_u^2}{J_r - (1 - m_u)^2} = \frac{1}{1/\mathbf{p}_r^T \mathbf{R}_r^{-1} \mathbf{p}_r - 1} \quad (3.114)$$

3.7.3 Simulation of LCL-MMSE and MMSE Receivers.

We refer to the LCL-MMSE receiver defined by Figure 3.10 as type I and the LCL-MMSE receiver defined by Figure 3.11 as type II. In this section we will demonstrate experimentally that the type I and type II LCL-MMSE receivers are indeed equivalent and furthermore compare the performances of the LCL-MMSE, MMSE, and real MMSE receivers.

The simulated system was a $K = 5$ user system with $N = 15$ chips per bit Gold sequences. No oversampling was done, and perfect power control was assumed, i.e., $Q = 1$ and $\gamma_1 = \gamma_2 = \dots = \gamma_K$. The propagation delays and phases were (randomly) selected to be

$$\boldsymbol{\tau}/T_c = [0 \quad 2.1942 \quad 7.7568 \quad 3.8888 \quad 3.8003]^T \quad (3.115)$$

$$\boldsymbol{\theta} = [0.7251 \quad 2.8904 \quad 3.9915 \quad 6.1891 \quad 1.7940]^T \quad (3.116)$$

For each Monte-Carlo run, independent realizations of the data and noise waveforms were created. The resulting realization of the received vector sequence was processed by the receivers, and the MSE and number of errors were recorded. The error probability was computed by the exact expression (3.8) and by the Gaussian approximation formula (3.20). The results of 10^6 Monte-Carlo runs are summarized in Table 3.1.

It is evident that the LCL-MMSE type I and II receivers are equivalent. Furthermore, there is a fair agreement between the simulated and theoretical quantities, and the Gaussian approximation is shown to be reasonably correct. These results indicate that the LCL-MMSE receiver performs better than the real MMSE receiver, which, in turn, is slightly better than the MMSE receiver.

Plots of the error probability as a function of the SNR, $2E_b/N_0$, is shown in Figure 3.13. The error probability for the different receiver were computed in three

Table 3.1: Results of simulation for SNR $2E_b/N_0 = 10$ dB.

		MMSE	Real	LCL-MMSE I	LCL-MMSE II
J	Theor.	0.2417	0.1297	0.1119	0.1119
	Simul.	0.2416	0.1297	0.1119	0.1119
P_e	Theor.	5.3385×10^{-3}	4.6819×10^{-3}	2.4139×10^{-3}	2.4139×10^{-3}
	Gauss.	5.4001×10^{-3}	4.7875×10^{-3}	2.4234×10^{-3}	2.4234×10^{-3}
	Simul.	5.3410×10^{-3}	4.6890×10^{-3}	2.3300×10^{-3}	2.3300×10^{-3}

ways: analytically, (o) joined by solid lines; by the Gaussian approximation, (x) joined by dashed lines; and by simulation, (*) joined by dotted lines. The curves are from the top: the MMSE receiver, the real MMSE receiver, and the LCL-MMSE I and II receivers (overlaid). For reference, the single-user system error probability, i.e., $Q(\sqrt{2E_b/N_0})$, is plotted as the bottommost curve. As seen from Figure 3.13, the LCL-MMSE receiver performs better than the real MMSE receiver, which, in turn, outperforms the MMSE receiver. These results can heuristically be explained as follows.

Recall that the error signal is formed as the difference between the desired user's bit and the filter output, i.e., $e(m) = d_1(m) - y_1(m)$. Note that only the real part of the error signal can cause the receiver to make a wrong decision, since $\hat{d}_1(m) = \text{sgn}[\text{Re}\{y_1(m)\}]$. Moreover, the complex MMSE receiver minimizes the *magnitude* of the error, and the real MMSE receiver minimizes the *real part* of the error. Thus, the complex receiver will generally perform worse than the real receiver.

The performance difference between the real MMSE and LCL-MMSE receivers can be explained by the following example. Suppose the desired user's phase is zero, i.e., $\theta_1 = 0$. This means that the imaginary branch of the type II LCL-MMSE receiver will not contain any information about the desired user. Thus, we are tempted

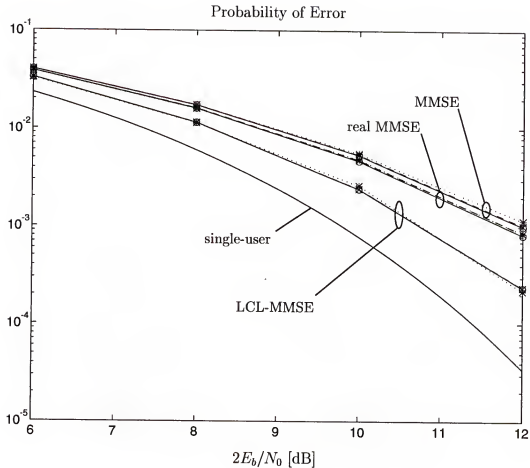


Figure 3.13: Error probability as a function of SNR

to delete the imaginary branch altogether. This will be a mistake, however, since plenty of information regarding the *interference* still remains in this branch. Thus, by processing this information and adding it to the soft decision, we can make the receiver more reliable. We can view the LCL-MMSE receiver as a form of *interference cancellation* on top of a linear receiver.

3.8 Conclusions

The main contributions of this chapter are found in Sections 3.5 and 3.6.

In Section 3.5, the basic properties of the MMSE receiver were derived. It was

shown that the MMSE receiver degenerates to the matched filter, if the multiple-access interference disappear, and to the decorrelating solution, if the noise disappear. This means that the MMSE receiver will be *the optimum receiver* in these two extreme situations (recall that optimum receiver here is defined as the linear single-user single-bit receiver with the smallest error probability).

In Section 3.6, complexity reduction through dimension reduction was proposed. The optimum form was derived and shown to be dependent on the unknown $\mathbf{A}$ matrix (recall that $\mathbf{A}$ is defined by the K users' code waveforms and propagation delays). The conclusion is therefore that we cannot implement the optimum transform in practice. However, several practical suboptimum schemes (the gamma, modified gamma, and symmetric dimension reduction (SDR) methods) and one asymptotically optimum scheme (the eigenvector identification (EGID) method) were proposed. These schemes were compared with the cyclically shifted filter bank (CSFB) which was the best previously known method [25].

The EGID scheme is clearly the best but also the most complex. The SDR is the simplest scheme and works slightly better than the CSFB method. The gamma filters did not perform as well as the other schemes, except when the interference is extremely large. Thus, the SDR is the method of choice if complexity is a concern. We should not use the EGID scheme if complexity is an issue. However, if the EGID receiver is allowed to observe the received signal for an extended period of time, it will (asymptotically) have the same performance as the full-length receiver in the sense that the attainable minimum mean-squared errors are the same. Moreover, since the excess mean-squared error will be less for the EGID scheme than for the full-length receiver, the EGID receiver might still be of practical interest.

In Section 3.7, a modification to the MMSE receiver, due to Aue and Reed [3], was presented. This modification has lower complexity and better performance than the ordinary MMSE receiver. It is believed that the results from Sections 3.5 and 3.6

are valid the modified receiver as well. This, however, is not proven at the present time and is subject to future work.

CHAPTER 4

PARAMETER ESTIMATION FOR THE AWGN CHANNEL

Many near-far resistant receivers have been proposed in the last decade, [57, 61, 22, 56], to mention just a few. All these schemes require exact knowledge of one or several parameters, such as each user's propagation delay, received power level, and carrier phase. In a practical system, these parameters need to be estimated and will therefore be subject to estimation errors. Relatively little has been done in terms of investigating the impact of such errors on receivers that assume perfect side-information and on how the estimation algorithms should be designed to minimize these adverse effects. However, lately there has been some work done in the area of *joint detection*—i.e., joint *data detection* and *parameter estimation*. See, for instance, Iltis, who uses an extended Kalman filter to obtain parameter estimates in a single-user scenario [18], and Radović and Aazhang, who combine the *expectation maximization* (EM) algorithm and the multistage detector in a near-far multiuser environment [37].

The standard method for *code acquisition*—i.e., initial propagation delay estimation, typically before data transmission—is the *sliding correlator* and modifications thereof [33, 38]. These methods are essentially single-user algorithms; they work reasonably well in a multiuser environment if the received powers are similar, but fail in a near-far environment [29, 28]. Other approaches to code acquisition are the *rapid acquisition by sequential estimation* (RASE) algorithm [59] and its multiuser extension the *multiple-access RASE* (MARASE) algorithm [30]. The sliding correlator, RASE, and MARASE algorithms assume that a known bit sequence is transmitted,

and thus they cannot be used for *tracking*—propagation delay estimation during data transmission.

The main contribution of this chapter is the treatment of the full estimation problem, i.e., estimation of the propagation delay, phase, and amplitude for all users in a DS-CDMA system. Most of the algorithms proposed, however, deal with propagation delay estimation. The reason for this is that obtaining accurate propagation delay estimates is a difficult problem, and there exist many algorithms for (suboptimally) estimating the remaining parameters, once reliable estimates of the propagation delays are available [34, 60, 9, 45]. No knowledge of the data sequences are required by the proposed algorithms. This means that the algorithms are applicable to both the acquisition and tracking problems.

The original work of this chapter is (1) the derivation of the Cramér-Rao bound, (2) the formulation of the maximum-likelihood, approximative maximum-likelihood, correlator, and MUSIC estimators, (3) the formulation of efficient search techniques of the MUSIC and correlator cost functions, and (4) the derivation of the optimum dimension reduction transform.

4.1 Cramér-Rao Bound

Since the derivation of the *Cramér-Rao bound* (CRB) is somewhat lengthy, we defer the details to Appendix A.

The parameter estimation problem can be formulated as follows: given M observations of the received vector, estimate the parameter vector $\phi \in \mathbb{R}^{3K+1}$,

$$\phi = [\sigma^2 \quad \gamma^T \quad \theta^T \quad \tau^T]^T \quad (4.1)$$

where σ^2 , γ , θ , and τ are defined by (2.4), (2.21)–(2.23). We will, however, concentrate our efforts on estimating the propagation delay vector τ . In fact, we will not discuss noise variance estimation at all. Nevertheless, we will still assume that the

elements of ϕ are unknown and deterministic parameters that are fixed during the interval for which we observe the received vector.

The CRB is a lower bound on the covariance matrix of any unbiased estimator [40]. Suppose $\hat{\phi}$ is an unbiased estimator of a vector of deterministic unknown parameters ϕ (i.e., $E[\hat{\phi}] = \phi$), then the covariance matrix of the estimator satisfies¹

$$E[(\hat{\phi} - \phi)(\hat{\phi} - \phi)^T] \geq \mathbf{J}^{-1} = \left(E \left[\left(\frac{\partial \ln L(\mathbf{r})}{\partial \phi} \right) \left(\frac{\partial \ln L(\mathbf{r})}{\partial \phi} \right)^T \right] \right)^{-1} \quad (4.2)$$

where $\mathbf{J} \in \mathbb{R}^{(1+3K) \times (1+3K)}$ is the *Fisher information matrix* and $\ln L(\mathbf{r})$ is the *log-likelihood function* of the observed data $\mathbf{r}$ with respect to ϕ .

In our case, the observed data is the received vector $\mathbf{r}(m)$ for $m = 1, 2, \dots, M$; hence, $\mathbf{r} = [\mathbf{r}^T(1) \ \dots \ \mathbf{r}^T(M)]^T \in \mathbb{C}^{MQN}$. In order to make the problem tractable, we condition the likelihood function on the transmitted bits, or, equivalently, on $\mathbf{z}(m)$. The resulting CRB may be interpreted as the CRB on estimators that are given knowledge of the transmitted bits. Nevertheless, the CRB presented here will still be a valid (although maybe not a tight) bound on the performance of estimators based on the model (2.17).

Since the noise is complex Gaussian, the conditional log-likelihood function (ignoring terms that are independent of ϕ) is

$$\ln L(\mathbf{r}) = -MQN \ln \sigma^2 - \frac{1}{\sigma^2} \sum_{m=1}^M \|\mathbf{r}(m) - \mathbf{A}\mathbf{B}\mathbf{z}(m)\|^2 \quad (4.3)$$

In order to calculate $\mathbf{J}$, the partial derivative with respect to the parameter vector ϕ is needed. As explained in Appendix A, it can be seen from (2.12) and (2.13) that $\partial/\partial\tau_k \mathbf{A}$ might not exist when $\delta_k = 0$. Therefore, in order to proceed, we will assume that $\delta_k \neq 0$ for $k = 1, \dots, K$. This rather mild restriction means that we are not sampling in perfect synchronization with any of the K users; furthermore, the algorithms proposed below are not dependent on this assumption.

¹With $\mathbf{X} \geq \mathbf{Y}$ for two symmetric positive semidefinite matrices $\mathbf{X}$ and $\mathbf{Y}$, we mean that $\mathbf{X} - \mathbf{Y}$ is symmetric positive semidefinite.

It is shown in Section A.1 that the information matrix may be written as

$$\mathbf{J} = \begin{bmatrix} MQN/\sigma^4 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\gamma\gamma} & \mathbf{J}_{\gamma\theta} & \mathbf{J}_{\gamma\tau} \\ \mathbf{0} & \mathbf{J}_{\gamma\theta}^T & \mathbf{J}_{\theta\theta} & \mathbf{J}_{\theta\tau} \\ \mathbf{0} & \mathbf{J}_{\gamma\tau}^T & \mathbf{J}_{\theta\tau}^T & \mathbf{J}_{\tau\tau} \end{bmatrix} \quad (4.4)$$

where the matrices $\mathbf{J}_{\gamma\gamma}$, $\mathbf{J}_{\gamma\theta}$, $\mathbf{J}_{\gamma\tau}$, $\mathbf{J}_{\theta\theta}$, $\mathbf{J}_{\theta\tau}$, $\mathbf{J}_{\tau\tau} \in \mathbb{R}^{K \times K}$ are defined as

$$\mathbf{J}_{\gamma\gamma} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \{ \mathbf{Z}^*(m) \mathbf{B}_\gamma^* \mathbf{A}^* \mathbf{A} \mathbf{B}_\gamma \mathbf{Z}(m) \} \quad (4.5)$$

$$\mathbf{J}_{\gamma\theta} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \{ \mathbf{Z}^*(m) \mathbf{B}_\gamma^* \mathbf{A}^* \mathbf{A} \mathbf{B}_\theta \mathbf{Z}(m) \} \quad (4.6)$$

$$\mathbf{J}_{\gamma\tau} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \{ \mathbf{Z}^*(m) \mathbf{B}_\gamma^* \mathbf{A}^* \Psi \mathbf{B} \mathbf{Z}(m) \} \quad (4.7)$$

$$\mathbf{J}_{\theta\theta} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \{ \mathbf{Z}^*(m) \mathbf{B}_\theta^* \mathbf{A}^* \mathbf{A} \mathbf{B}_\theta \mathbf{Z}(m) \} \quad (4.8)$$

$$\mathbf{J}_{\theta\tau} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \{ \mathbf{Z}^*(m) \mathbf{B}_\theta^* \mathbf{A}^* \Psi \mathbf{B} \mathbf{Z}(m) \} \quad (4.9)$$

$$\mathbf{J}_{\tau\tau} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \{ \mathbf{Z}^*(m) \mathbf{B}^* \Psi^* \Psi \mathbf{B} \mathbf{Z}(m) \} \quad (4.10)$$

where $\mathbf{B}_\theta = \mathbf{j} \mathbf{B}$,

$$\mathbf{B}_\gamma = \text{diag}(\exp(j\theta_1), \exp(j\theta_1), \dots, \exp(j\theta_K), \exp(j\theta_K)) \in \mathbb{C}^{2K \times 2K} \quad (4.11)$$

and where $\mathbf{Z}(m) \in \{-1, 0, 1\}^{2K \times K}$ is defined as

$$(\mathbf{Z}(m))(i, k) = \begin{cases} z_{2k-1}(m) & \text{if } i = 2k-1 \\ z_{2k}(m) & \text{if } i = 2k \\ 0 & \text{otherwise} \end{cases} \quad (4.12)$$

and $\Psi = [\psi_1 \ \psi_2 \ \dots \ \psi_{2K}] \in \mathbb{R}^{Q \times 2K}$ is defined as

$$\psi_{2k-1} = \frac{\partial \mathbf{a}_{2k-1}}{\partial \tau_k} = \frac{1}{T_i} [\mathbf{D}(p_k + 1, 1) - \mathbf{D}(p_k, 1)] \mathbf{c}_k \quad (4.13)$$

$$\psi_{2k} = \frac{\partial \mathbf{a}_{2k}}{\partial \tau_k} = \frac{1}{T_i} [\mathbf{D}(p_k + 1, -1) - \mathbf{D}(p_k, -1)] \mathbf{c}_k \quad (4.14)$$

It is also shown in Section A.1 that the *asymptotic Fisher information matrix*, $\mathbf{J}_\infty$, is

$$\mathbf{J}_\infty = \lim_{M \rightarrow \infty} \mathbf{J} = \begin{bmatrix} MQN/\sigma^4 & 0 & 0 & 0 \\ 0 & \mathbf{J}_{\gamma\gamma,\infty} & 0 & \mathbf{J}_{\gamma\tau,\infty} \\ 0 & 0 & \mathbf{J}_{\theta\theta,\infty} & 0 \\ 0 & \mathbf{J}_{\gamma\tau,\infty}^T & 0 & \mathbf{J}_{\tau\tau,\infty} \end{bmatrix} \quad (4.15)$$

where $\mathbf{J}_{\gamma\gamma,\infty}, \mathbf{J}_{\gamma\tau,\infty}, \mathbf{J}_{\theta\theta,\infty}, \mathbf{J}_{\tau\tau,\infty} \in \mathbb{R}^{K \times K}$ are

$$\mathbf{J}_{\gamma\gamma,\infty} = \frac{M}{\sigma^2} \mathbf{G}^* \text{diag}(\mathbf{A}^* \mathbf{A}) \mathbf{G} \quad (4.16)$$

$$\mathbf{J}_{\gamma\tau,\infty} = \frac{M}{\sigma^2} \mathbf{G} \text{diag}(\mathbf{A}^* \Psi) \mathbf{G} \sqrt{\mathbf{P}} \quad (4.17)$$

$$\mathbf{J}_{\theta\theta,\infty} = \frac{M}{\sigma^2} \mathbf{G} \text{diag}(\mathbf{A}^* \mathbf{A}) \mathbf{G} \mathbf{P} \quad (4.18)$$

$$\mathbf{J}_{\tau\tau,\infty} = \frac{M}{\sigma^2} \mathbf{G} \text{diag}(\Psi^* \Psi) \mathbf{G} \mathbf{P} \quad (4.19)$$

where $\mathbf{G} \in \{0, 1\}^{2K \times K}$ and $\mathbf{P} \in \mathbb{R}^{K \times K}$ are defined as

$$\mathbf{G}(i, k) = \begin{cases} 1 & \text{if } i = 2k - 1 \\ 1 & \text{if } i = 2k \\ 0 & \text{otherwise} \end{cases} \quad (4.20)$$

$$\mathbf{P} = \text{diag}(P_1, P_2, \dots, P_K) \quad (4.21)$$

It is possible to invert the asymptotic information matrix analytically, since all blocks are diagonal. Furthermore, the structure of $\mathbf{J}_\infty$ is such that the bound on the variance of an unbiased estimator of γ_1 , θ_1 , or τ_1 is independent of the parameters for any other user. In particular, the bound on the variance of a τ_1 -estimator is independent of P_k for $k = 2, 3, \dots, K$. This is important since it gives us hope that there indeed may exist estimators that are near-far resistant. Practical advantages with the asymptotic expression (4.15) are (1) it is independent of the realization of $\mathbf{z}(m)$ and (2) it is (much) less expensive to compute than the finite sample expres-

sion (4.4). Furthermore, as indicated in Section 4.2.5, the finite sample information matrix approaches the asymptotic expression for rather modest values of M .

4.2 Estimation Algorithms

4.2.1 Maximum-Likelihood

The *maximum-likelihood* (ML) estimator is found by maximizing the log-likelihood function (4.3) with respect to ϕ . As seen from (4.3), the optimization problem is separable with respect to σ^2 . In particular, we can find the ML estimates of γ , θ , and τ independently of σ^2 .

Maximizing the log-likelihood function is equivalent to minimizing the negative log-likelihood function given by (constants unaffected by the optimization procedure are ignored)

$$\ell(\gamma, \theta, \tau) = \sum_{m=1}^M \|\mathbf{r}(m) - \mathbf{A}\mathbf{B}\mathbf{z}(m)\|^2 \quad (4.22)$$

Note that the log-likelihood function is conditioned on $\mathbf{z}(m)$ for $m = 1, 2, \dots, M$. Thus, minimizing (4.22) with respect to ϕ requires knowledge of the vector $\mathbf{z} \in \{-1, 0, 1\}^{2MK}$ defined as

$$\mathbf{z} = [\mathbf{z}^T(1) \quad \mathbf{z}^T(2) \quad \cdots \quad \mathbf{z}^T(M)]^T \quad (4.23)$$

Clearly, $\mathbf{z}$ is unknown, and we are forced to detect $\mathbf{z}$ from the observed vectors. Thus, the maximum likelihood estimates $\hat{\gamma}^{\text{ML}}$, $\hat{\theta}^{\text{ML}}$, $\hat{\tau}^{\text{ML}}$, and $\hat{\mathbf{z}}^{\text{ML}}$ are found as

$$\begin{bmatrix} \hat{\gamma}^{\text{ML}} \\ \hat{\theta}^{\text{ML}} \\ \hat{\tau}^{\text{ML}} \\ \hat{\mathbf{z}}^{\text{ML}} \end{bmatrix} = \arg \min_{\substack{\gamma \\ \theta \\ \tau \\ \mathbf{z}}} \sum_{m=1}^M \|\mathbf{r}(m) - \mathbf{A}\mathbf{B}\mathbf{z}(m)\|^2 \quad (4.24)$$

Note that the maximum-likelihood algorithm estimates γ , τ , θ , and $\mathbf{z}$ simultaneously. This estimator exploits the full structure of the problem and will achieve the CRB [40].

The optimization of the log-likelihood function over $\mathbf{z}$ as well as over ϕ is unfortunate, since we are facing an optimization problem with both continuous and discrete parameters. The standard algorithms for optimization, e.g., Gauss-Newton or Levenberg-Marquardt, are not designed to work with discrete parameters. Nevertheless, these algorithms can still be used to find the maximum-likelihood estimate of ϕ as follows. Substitute a realization of $\mathbf{z}$ into the log-likelihood function, minimize the log-likelihood function with respect to the continuous parameters, save the parameters and the value of the cost function, and repeat this procedure until all possible realization of $\mathbf{z}$ have been tested. The ML estimates are then found as the parameters corresponding to the smallest value of the cost function. The problem is thus solvable in principle, but the complexity becomes overwhelming as the number of users or the observation interval grows.

4.2.2 Approximative Maximum Likelihood

The ML estimator is impractical since we need to optimize a cost function with respect to continuous as well as discrete parameters. However, we can make an approximation that in essence allows us to “forget” that we have a mixed optimization problem. If we define

$$\mathbf{s}(m) = \mathbf{B}\mathbf{z}(m) \in \mathbb{C}^{QN} \quad (4.25)$$

then the received vector may be written as

$$\mathbf{r}(m) = \mathbf{A}\mathbf{s}(m) + \mathbf{n}(m) \quad (4.26)$$

If we assume that $\mathbf{s}(m)$ consists of continuous, unknown and deterministic parameters, then the ML-estimate of $\mathbf{s}(m)$ is $\hat{\mathbf{A}}^\dagger \mathbf{r}(m)$, where the matrix $\hat{\mathbf{A}}^\dagger$ is the left pseudoinverse of $\hat{\mathbf{A}} = \mathbf{A}(\hat{\boldsymbol{\tau}})$, [13, 52]. Since $\hat{\mathbf{A}}$ has full rank, we can compute $\hat{\mathbf{A}}^\dagger$ as

$$\hat{\mathbf{A}}^\dagger = (\hat{\mathbf{A}}^* \hat{\mathbf{A}})^{-1} \hat{\mathbf{A}}^* \quad (4.27)$$

The *approximative maximum-likelihood* (AML) estimate of τ is formulated as

$$\begin{aligned}\hat{\tau}^{\text{AML}} &= \arg \min_{\tau} \sum_{m=1}^M \left\| \mathbf{r}(m) - \mathbf{A} \mathbf{B} \mathbf{z}(m) \right\|^2 \Big|_{\mathbf{B} \mathbf{z}(m) = \mathbf{A}^\dagger(\tau) \mathbf{r}(m)} \\ &= \arg \min_{\tau} \text{trace} \left(\mathbf{P}_{\mathbf{A}(\tau)}^\perp \hat{\mathbf{R}}_M \right)\end{aligned}\quad (4.28)$$

where $\mathbf{P}_{\mathbf{A}(\tau)}^\perp = \mathbf{I}_{QN} - \mathbf{A}(\tau) \mathbf{A}^\dagger(\tau)$ is the *orthogonal projection matrix* onto the subspace $[\text{range}(\mathbf{A}(\tau))]^\perp$, and $\hat{\mathbf{R}}_M$ is the sample correlation matrix

$$\hat{\mathbf{R}}_M = \frac{1}{M} \sum_{m=1}^M \mathbf{r}(m) \mathbf{r}^*(m) \quad (4.29)$$

Note that in this procedure, we are not using the full structure of $\mathbf{s}(m)$ indicated by (4.25). We therefore expect (4.28) to yield worse estimates than (4.24). Also, note that the AML algorithm estimates all the delays simultaneously.

4.2.3 MUSIC

The foundation to the *multiple signal classification* (MUSIC) algorithm [41], originally formulated for array processing, is the observation that the columns of $\mathbf{A}$ are contained in the signal subspace, $\text{range}(\mathbf{E}_s)$, or, equivalently, are orthogonal to all vectors in the noise subspace $\text{range}(\mathbf{E}_n)$.

Consider the vectors $\mathbf{b}_{k,1}(\tau)$ and $\mathbf{b}_{k,2}(\tau)$,

$$\mathbf{b}_{k,1}(\tau) = \left[\frac{\delta}{T_i} \mathbf{D}(p+1, 1) + \left(1 - \frac{\delta}{T_i} \right) \mathbf{D}(p, 1) \right] \mathbf{c}_k \quad (4.30)$$

$$\mathbf{b}_{k,2}(\tau) = \left[\frac{\delta}{T_i} \mathbf{D}(p+1, -1) + \left(1 - \frac{\delta}{T_i} \right) \mathbf{D}(p, -1) \right] \mathbf{c}_k \quad (4.31)$$

where $\tau = pT_i + \delta$ is such that p is an integer and $\delta = [0, T_i)$. We observe from (2.12) and (2.13) that $\mathbf{a}_{2k-1} = \mathbf{b}_{k,1}(\tau_k)$ and $\mathbf{a}_{2k} = \mathbf{b}_{k,2}(\tau_k)$. Thus, given knowledge of $\mathbf{R}$ we can find τ_k as the solution to $\|\mathbf{E}_n^* \mathbf{b}_{k,1}(\tau)\|^2 = 0$ or as the solution to $\|\mathbf{E}_n^* \mathbf{b}_{k,2}(\tau)\|^2 = 0$. Moreover, the solution is unique under the assumptions that $2K < QN$ and that $\mathbf{A}$ has full rank for all possible realizations of τ . In practice, the correlation matrix is

unknown and is therefore estimated by the sample correlation matrix. A consistent estimate of $\mathbf{E}_n$ is found in the eigenvalue decomposition of $\hat{\mathbf{R}}_M$,

$$\hat{\mathbf{R}}_M = \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s \hat{\mathbf{E}}_s^* + \hat{\mathbf{E}}_n \hat{\mathbf{\Lambda}}_n \hat{\mathbf{E}}_n^* \quad (4.32)$$

where the columns of $\hat{\mathbf{E}}_n$ are the eigenvectors corresponding to the $QN - 2K$ smallest eigenvalues of $\hat{\mathbf{R}}_M$. Note, however, that the columns of $\mathbf{A}$ will now be only approximately orthogonal to the columns of $\hat{\mathbf{E}}_n$.

The original MUSIC algorithm is formulated for the case where each column in $\mathbf{A}$ is a function of a distinct parameter. However, that is not the case here. One possible modification of the MUSIC algorithm is as follows. Let the MUSIC estimate of τ_k be defined as

$$\hat{\tau}_k^{\text{MU}} = \arg \min_{\tau \in [0, T)} J_{k, \text{MU}}(\tau) \quad (4.33)$$

where $J_{k, \text{MU}}(\tau)$ is the *MUSIC cost function* for the k th user defined as

$$J_{k, \text{MU}}(\tau) = \frac{\|\hat{\mathbf{E}}_n^* \mathbf{b}_{k,1}(\tau)\|^2}{\|\mathbf{b}_{k,1}(\tau)\|^2} + \frac{\|\hat{\mathbf{E}}_n^* \mathbf{b}_{k,2}(\tau)\|^2}{\|\mathbf{b}_{k,2}(\tau)\|^2} \quad (4.34)$$

As seen from (4.30) and (4.31), $\mathbf{b}_{k,1}(\tau)$ and $\mathbf{b}_{k,2}(\tau)$ are piecewise linear in τ . In particular, for $\tau \in [pT_i, (p+1)T_i)$, referred to as the p th bin of τ ,

$$\mathbf{b}_{k,1}(\tau) = \frac{\delta}{T_i} \mathbf{b}_{k,1}([p+1]T_i) + \left(1 - \frac{\delta}{T_i}\right) \mathbf{b}_{k,1}(pT_i) \quad (4.35)$$

$$\mathbf{b}_{k,2}(\tau) = \frac{\delta}{T_i} \mathbf{b}_{k,2}([p+1]T_i) + \left(1 - \frac{\delta}{T_i}\right) \mathbf{b}_{k,2}(pT_i) \quad (4.36)$$

By substituting (4.35) and (4.36) into (4.34) and noting that if p is an integer then $\|\mathbf{b}_{k,1}(pT_i)\|^2 = \|\mathbf{b}_{k,2}(pT_i)\|^2 = QN$, we obtain that for $\tau \in [pT_i, (p+1)T_i)$,

$$\begin{aligned} J_{k, \text{MU}}(\tau) &= \frac{s^2 K_{11} + (1-s)^2 K_{12} + 2s(1-s) K_{13}}{(s^2 + (1-s)^2) QN + 2s(1-s) K_{14}} \\ &\quad + \frac{s^2 K_{21} + (1-s)^2 K_{22} + 2s(1-s) K_{23}}{(s^2 + (1-s)^2) QN + 2s(1-s) K_{24}} \end{aligned} \quad (4.37)$$

where we have defined $s = \delta/T_i$ and

$$K_{11} = \|\hat{\mathbf{E}}_n^* \mathbf{b}_{k,1}([p+1]T_i)\|^2 \quad (4.38)$$

$$K_{12} = \|\hat{\mathbf{E}}_n^* \mathbf{b}_{k,1}(pT_i)\|^2 \quad (4.39)$$

$$K_{13} = \text{Re}\{\mathbf{b}_{k,1}^*([p+1]T_i)\hat{\mathbf{E}}_n\hat{\mathbf{E}}_n^*\mathbf{b}_{k,1}(pT_i)\} \quad (4.40)$$

$$K_{14} = \mathbf{b}_{k,1}^*([p+1]T_i)\mathbf{b}_{k,1}(pT_i) \quad (4.41)$$

and where K_{21} , K_{22} , K_{23} , and K_{24} are defined in an analogous manner by changing $\mathbf{b}_{k,1}$ to $\mathbf{b}_{k,2}$ in (4.38)–(4.41).

From (4.37), we see that $J_{k,\text{MU}}(\tau)$ can be written as a rational function of two polynomials of degree four, and furthermore that $J_{k,\text{MU}}(\tau)$ is differentiable for $\tau \in [0, T)$, except when $\tau = pT_i$ for an integer p . This suggests an efficient search procedure of the cost function. We form the set $\mathcal{T}$ of candidate MUSIC estimates as the union of the set $\{0, T_i, 2T_i, \dots, (QN-1)T_i\}$ and the set of real solutions to the equation

$$\frac{dJ_{k,\text{MU}}(\tau)}{d\tau} = 0 \quad (4.42)$$

The solutions to (4.42), i.e., the roots to the numerator polynomial of (4.42), can be found efficiently by numerical methods. Finally, $\hat{\tau}_k^{\text{MU}}$ is found as the member of $\mathcal{T}$ corresponding to the smallest value of the cost function.

Note that the MUSIC algorithm finds one element of τ at a time.

4.2.4 Sliding Correlator

The standard approach for propagation delay estimation is the so-called *sliding correlator* [33]. Much as the standard receiver (i.e., the matched filter), the sliding correlator is only optimum in a single-user situation or when the received code waveforms are orthogonal. It is therefore expected that this scheme will fail in a near-far situation. However, for the purpose of comparison, we will now describe an algorithm that is analogous to the sliding correlator.

If we constrain $d_k(m) = 1$ for $m = 1, \dots, M$, we see from (2.8) that the contribution from the k th user to the received vector is $\beta_k \mathbf{a}_{2k-1}$. Thus, since $\mathbf{a}_{2k-1} = \mathbf{b}_{k,1}(\tau_k)$ is roughly orthogonal to $\{\mathbf{a}_{2j-1}, \mathbf{a}_{2j}\}$ for $j \neq k$, and since $\{\mathbf{a}_{2j-1}, \mathbf{a}_{2j}\}$ are independent of τ_k for $j \neq k$, we can form the correlator estimate of τ_k as

$$\hat{\tau}_k^c = \arg \min_{\tau_k \in [0, T)} J_{k,c}(\tau) \quad (4.43)$$

where $J_{k,c}(\tau)$ is the *sliding correlator cost function* for the k th user

$$J_{k,c}(\tau) = - \left| \frac{1}{M} \sum_{m=1}^M \frac{\mathbf{b}_{k,1}^*(\tau) \mathbf{r}(m)}{\|\hat{\mathbf{b}}_{k,1}(\tau)\|} \right|^2 = - \frac{|\mathbf{b}_{k,1}^*(\tau) \mathbf{r}_M|^2}{\|\hat{\mathbf{b}}_{k,1}(\tau)\|^2} \quad (4.44)$$

and where $\mathbf{r}_M = \frac{1}{M} \sum_{m=1}^M \mathbf{r}(m)$. As seen from this equation, the correlator, as the MUSIC algorithm, estimates one propagation delay at a time.

We now form two equations analogous to (4.37) and (4.42). For $\tau \in [pT_i, (p+1)T_i)$, $J_{k,c}(\tau)$ can be written as

$$J_{k,c}(\tau) = - \frac{s^2 |C_1 - C_0|^2 + 2s \operatorname{Re}\{(C_1 - C_0)C_0^*\} + |C_0|^2}{(s^2 + (1-s)^2)QN + 2s(1-s)\mathbf{b}_{k,1}^*([p+1]T_i)\mathbf{a}_{k,1}(pT_i)} \quad (4.45)$$

where $C_1 = \mathbf{b}_{k,1}^*([p+1]T_i)\mathbf{r}_M$ and $C_0 = \mathbf{b}_{k,1}^*(pT_i)\mathbf{r}_M$. After differentiating (4.45) with respect to τ , we note the resulting numerator is of degree two. The solution to $dJ_{k,c}(\tau)/d\tau = 0$ can therefore be found analytically, and the search for a global minimum can be carried out in the same manner as in Section 4.2.3.

Note that the MUSIC, ML, and AML algorithms do not assume that $d_k(m) = 1$. Therefore, in the simulations reported in Section 4.2.5, the correlator is given some extra side-information, and the comparisons with the other algorithms are actually biased in favor of the correlator.

4.2.5 Numerical Results

In this section, the performance of the AML, MUSIC, and correlator algorithms is evaluated by Monte-Carlo simulations.

The simulated system was a 10-user system with $N = 31$ chips per bit Gold sequences, generated by the polynomials $g_1(x) = x^5 + x + 1$ and $g_2(x) = x^5 + x^4 + x^3 + x^2 + 1$, [39]. Oversampling was not used, $Q = 1$, and the signal-to-noise ratio of the first user was fixed at $\text{SNR}_1 = 2E_{b,1}/N_0 = 15$ dB. The *near-far ratio* is defined as P_2/P_1 , where all interfering users have the same received power, $P_k = P_2$ for $k = 3, 4, \dots, 10$.

For each Monte-Carlo run, independent realizations of the noise and data sequences were created; however, the propagation delays, received amplitudes, and phases were fixed. The measure of performance for an estimator is the sample standard deviation, $\text{std}(\hat{\tau}_1)$,

$$\text{std}^2(\hat{\tau}_1) = \frac{1}{R} \sum_{m=1}^R \left(\hat{\tau}_1(m) - \frac{1}{R} \sum_{k=1}^R \hat{\tau}_1(k) \right)^2 \quad (4.46)$$

where $R = 500$ is the number of Monte-Carlo runs that were performed for each simulation. The estimators turned out to be approximately unbiased, and hence the sample standard deviation is an appropriate performance measure.

The absolute performance of the propagation delay estimators is of course dependent on the particular τ for which the simulations were done; however, the relative performance of the estimators proved to be roughly independent of the realization of the delays.

The Cramér-Rao bound serves as a lower bound on the standard deviation. The asymptotic CRB, given by (4.15), was used since the ordinary CRB is dependent on the data sequences, which were changed between Monte-Carlo runs. Moreover, the actual CRB was very close to the asymptotic expression if the observation interval M was greater than 20.

The optimization of the MUSIC and sliding correlator cost functions was carried out as described in Sections 4.2.3 and 4.2.4. This guarantees that the global minimum is found. The cost functions are highly nonlinear, and sometimes the estimates found

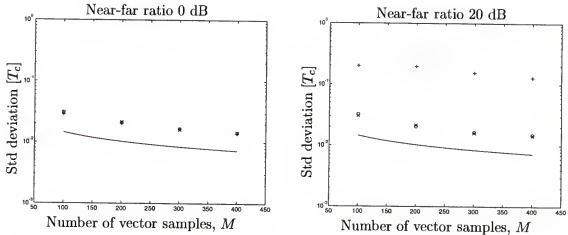


Figure 4.1: Standard deviation of the AML: (x), the MUSIC: (o), and the correlator estimators: (+). The solid line is the Cramér-Rao Bound.

were more than one chip duration away from the true delay. Such estimates are called *outliers*. The error due to an outlier estimate may be very large. Therefore, to make the comparison with the CRB meaningful, the outlier errors were removed before the standard deviation was computed

An outlier error is of course disastrous in the acquisition phase. However, in the tracking phase, we will probably limit the search of the cost function to be in the vicinity (say, plus minus $T_c/2$) of the previous estimate. This is reasonable since the propagation delays are typically slowly varying. By limiting the search, the algorithm will not find an outlier estimate.

The MUSIC estimator produced very few outliers when more than $M = 100$ vector samples were observed. For $M = 50$ and near-far ratio 20 dB, the relative number of outliers increased to about 20%. On the other hand, the sliding correlator had a high number of outliers in a near-far situation even if the observation interval was made long. When the near-far ratio was 20 dB, the outlier ratio was 50% for $M = 100$ and 13% for $M = 400$.

In case of the AML algorithm, the optimization was carried out by a numerical

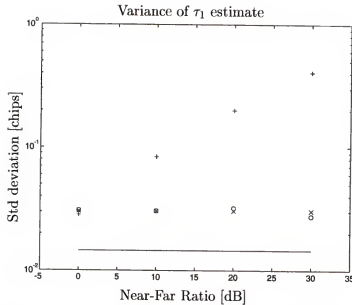


Figure 4.2: Standard deviation of the AML: (x), the MUSIC: (o), and the correlator estimators: (+). The solid line is the Cramér-Rao Bound, and the number of observed vector samples is $M = 100$.

search started in the true point. The cost function is such that a local minimum is often located very close to the true point. Therefore, there is a lower risk of finding an outlier error, and the performance of the AML algorithm shown here is therefore slightly optimistic.

In Figure 4.1, the standard deviations of the AML, MUSIC, and correlation algorithms are plotted versus the number of observed vectors, and in Figure 4.2, the standard deviations are plotted as functions of the near-far ratio. Note that the MUSIC and AML algorithms performed similarly and were unaffected by the near-far problem. However, the correlator algorithm was very sensitive to interference. Furthermore, none of the estimators attained the Cramér-Rao bound. Thus, the ML estimator would have had the best performance, but also largest complexity.

The results presented here were obtained with $\text{SNR}_1 = 15$ dB. The MUSIC estimator seems to be more noise sensitive than the correlator and AML algorithms.

With good power control and where the SNR is low, the correlator works as good as or better than the MUSIC and AML estimators. When SNR_1 is low and the near-far ratio is high, the AML is probably the method of choice. However, when SNR_1 is reasonably high and the interference significant, the MUSIC algorithm is attractive since it has a much lower complexity than the AML method.

4.3 Dimension Reduction

Many good estimators have a computational complexity on the order of $(QN)^3$, where QN is the dimensionality of the observed data. One possibility to lower the complexity of the algorithms is to reduce the dimensionality of the received vector by premultiplying with $\mathbf{V}^*$, where $\mathbf{V} \in \mathbb{C}^{QN \times L}$. Obviously, we would like to choose $\mathbf{V}$ such that it is optimum in some sense. If we restrict our attention to unbiased estimators, the natural choice of performance measure is the Cramér-Rao bound.

4.3.1 Optimum Dimension Reduction

Since we are mostly interested in estimating $\boldsymbol{\gamma}$, $\boldsymbol{\theta}$, and $\boldsymbol{\tau}$, we define the optimum transform $\mathbf{V} \in \mathbb{C}^{QN \times L}$ to be such that the CRB for these parameters is conserved. That is, $\mathbf{V}$ is an optimum transform if the CRB for an estimator based on the received vectors $\{\mathbf{r}(1), \dots, \mathbf{r}(M)\}$ is the same as the CRB for an estimator based on $\{\mathbf{V}^*\mathbf{r}(1), \dots, \mathbf{V}^*\mathbf{r}(M)\}$. Note that, by definition, no $\mathbf{V}$ exists such that the CRB is reduced. The idea of dimension reduction is not new, and the work presented in this section is inspired in many respects by Anderson [2].

The reduced dimension received vector, $\mathbf{r}_V(m) \in \mathbb{C}^L$, can be written as

$$\mathbf{r}_V(m) = \mathbf{V}^*\mathbf{r}(m) = \mathbf{V}^*\mathbf{A}\mathbf{B}\mathbf{z}(m) + \mathbf{V}^*\mathbf{n}(m) = \mathbf{A}_V\mathbf{B}\mathbf{z}(m) + \mathbf{n}_V(m) \quad (4.47)$$

where we have defined

$$\mathbf{A}_V = \mathbf{V}^*\mathbf{A} = [\mathbf{a}_{V,1} \quad \dots \quad \mathbf{a}_{V,2K}] \in \mathbb{C}^{L \times 2K} \quad (4.48)$$

$$\mathbf{n}_V(m) = \mathbf{V}^* \mathbf{n}(m) \quad (4.49)$$

We now make the assumption that $\mathbf{V}^* \mathbf{V} = \mathbf{I}_L$. Since $\mathbf{n}(m)$ is zero-mean complex Gaussian and has second moment matrices as given by (2.7), $\mathbf{n}_V(m) = \mathbf{V}^* \mathbf{n}(m)$ will also be zero-mean complex Gaussian with second moment matrices

$$\begin{aligned} \mathbb{E}[\mathbf{n}_V(m) \mathbf{n}_V^*(k)] &= \mathbf{V}^* \mathbb{E}[\mathbf{n}(m) \mathbf{n}^*(k)] \mathbf{V} = \mathbf{V}^* (\sigma^2 \mathbf{I}_N \delta_K(k - m)) \mathbf{V} \\ &= \sigma^2 \mathbf{I}_L \delta_K(k - m) \end{aligned} \quad (4.50)$$

$$\mathbb{E}[\mathbf{n}_V(m) \mathbf{n}_V^T(k)] = \mathbf{0} \quad (4.51)$$

Thus, the log-likelihood function (ignoring constants independent of ϕ) is

$$\ln L(\mathbf{r}_V) = -ML \ln \sigma^2 - \frac{1}{\sigma^2} \sum_{m=1}^M \|\mathbf{r}_V(m) - \mathbf{A}_V \mathbf{B} \mathbf{z}(m)\|^2 \quad (4.52)$$

Moreover, if we assume that the columns of $\mathbf{A}_V$ are linearly independent for all possible realizations of $\boldsymbol{\tau}$ and exclude the case when $\delta_k = 0$ for any $k = 1, 2, \dots, K$, then the Fisher information matrix for an unbiased estimator based on $\mathbf{r}_V$ will be as in (4.4) with the appropriate substitutions, $\mathbf{A} \leftarrow \mathbf{A}_V$, $\boldsymbol{\Psi} \leftarrow \boldsymbol{\Psi}_V = \mathbf{V}^* \boldsymbol{\Psi}$, and $QN \leftarrow L$,

$$\mathbf{J}_V = \begin{bmatrix} ML/\sigma^4 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{V,\gamma\gamma} & \mathbf{J}_{V,\gamma\theta} & \mathbf{J}_{V,\gamma\tau} \\ \mathbf{0} & \mathbf{J}_{V,\gamma\theta}^T & \mathbf{J}_{V,\theta\theta} & \mathbf{J}_{V,\theta\tau} \\ \mathbf{0} & \mathbf{J}_{V,\gamma\tau}^T & \mathbf{J}_{V,\theta\tau}^T & \mathbf{J}_{V,\tau\tau} \end{bmatrix} \quad (4.53)$$

where

$$\mathbf{J}_{V,\gamma\gamma} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \left\{ \mathbf{Z}^*(m) \mathbf{B}_\gamma^* \mathbf{A}_V^* \mathbf{A}_V \mathbf{B}_\gamma \mathbf{Z}(m) \right\} \quad (4.54)$$

$$\mathbf{J}_{V,\gamma\theta} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \left\{ \mathbf{Z}^*(m) \mathbf{B}_\gamma^* \mathbf{A}_V^* \mathbf{A}_V \mathbf{B}_\theta \mathbf{Z}(m) \right\} \quad (4.55)$$

$$\mathbf{J}_{V,\gamma\tau} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \left\{ \mathbf{Z}^*(m) \mathbf{B}_\gamma^* \mathbf{A}_V^* \boldsymbol{\Psi}_V \mathbf{B} \mathbf{Z}(m) \right\} \quad (4.56)$$

$$\mathbf{J}_{V,\theta\theta} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \{ \mathbf{Z}^*(m) \mathbf{B}_\theta \mathbf{A}_V^* \mathbf{A}_V \mathbf{B}_\theta \mathbf{Z}(m) \} \quad (4.57)$$

$$\mathbf{J}_{V,\theta\tau} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \{ \mathbf{Z}^*(m) \mathbf{B}_\theta^* \mathbf{A}_V^* \Psi_V \mathbf{B} \mathbf{Z}(m) \} \quad (4.58)$$

$$\mathbf{J}_{V,\tau\tau} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \{ \mathbf{Z}^*(m) \mathbf{B}^* \Psi_V^* \Psi_V \mathbf{B} \mathbf{Z}(m) \} \quad (4.59)$$

The form of the optimum dimension reduction transform is given in the following theorem.

THEOREM 4.1 *If $\mathbf{V} \in \mathbb{C}^{Q^N \times L}$ is such that $\mathbf{V}^* \mathbf{V} = \mathbf{I}_L$ and*

$$\text{range}([\mathbf{A} \quad \Psi]) \subseteq \text{range}(\mathbf{V}) \quad (4.60)$$

then the CRB for an estimator of the parameter vector $\phi' = [\gamma^T \quad \theta^T \quad \tau^T]^T$ is the same for an estimator based on the observed data $\{\mathbf{r}(1), \dots, \mathbf{r}(M)\}$ as for an estimator based on $\{\mathbf{V}^* \mathbf{r}(1), \dots, \mathbf{V}^* \mathbf{r}(M)\}$.

PROOF: The CRB with respect to ϕ' is given as the inverse of the lower right-hand $3K \times 3K$ block of the Fisher information matrix (4.53). To prove the theorem, it is therefore sufficient to show that this block of the information matrix is the same for an estimator based on $\{\mathbf{r}(1), \dots, \mathbf{r}(M)\}$ as for an estimator based on $\{\mathbf{V}^* \mathbf{r}(1), \dots, \mathbf{V}^* \mathbf{r}(M)\}$.

Since the columns of $\mathbf{V}$ form an orthonormal basis for $\mathbf{W} = \text{range}(\mathbf{V}) \subseteq \mathbb{C}^N$, the orthogonal projection matrix $\mathbf{P}_\mathbf{W}$ onto $\mathbf{W}$ is found as $\mathbf{P}_\mathbf{W} = \mathbf{V} \mathbf{V}^*$. By assumption, $\text{range}([\mathbf{A} \quad \Psi]) \subseteq \mathbf{W}$ and therefore

$$\mathbf{V} \mathbf{V}^* \mathbf{A} = \mathbf{A} \quad (4.61)$$

$$\mathbf{V} \mathbf{V}^* \Psi = \Psi \quad (4.62)$$

This implies that $\mathbf{A}_V^* \mathbf{A}_V = \mathbf{A}^* \mathbf{V} \mathbf{V}^* \mathbf{A} = \mathbf{A}^* \mathbf{A}$ and, in the same manner, that $\Psi_V^* \Psi_V = \Psi^* \Psi$ and $\mathbf{A}_V^* \Psi_V = \mathbf{A}^* \Psi$. It is now seen that the blocks of $\mathbf{J}_V$, (4.54)–(4.59), and $\mathbf{J}$, (4.5)–(4.10), are equivalent, and the theorem follows. $\blacksquare$

If we assume that $[\mathbf{A} \quad \mathbf{\Psi}]$ has full rank, the theorem says that we can lower the dimension of the received vectors from QN to $4K$ without loss of performance in the CRB sense.

It is easy to show that, as $M \rightarrow \infty$, the MUSIC estimates of τ_k based on $\{\mathbf{r}(1), \dots, \mathbf{r}(M)\}$ and $\{\mathbf{r}_V(1), \dots, \mathbf{r}_V(M)\}$, will be identical and equal to τ_k . This, however, is only a statement of consistency and does not imply that the estimator variance will decrease with M at the same speed in both cases. Nevertheless, in the next section, we will demonstrate experimentally that the optimum transform indeed has this desirable property.

4.3.2 Numerical Results

The optimum transform conserves the CRB which is a lower bound on the variance of an unbiased estimator. Thus, there is no guarantee that a suboptimum method will perform equally well before and after dimension reduction. However, in this section, we will verify this result for the MUSIC estimator by simulations.

The simulated system was a 2-user system with $Q = 1$, $N = 15$ chips per bit, $T_c = 1$, $\text{SNR}_1 = 2E_{b,1}/N_0 = 15$ dB, and $P_2 = P_1$ (i.e., perfect power control). The code sequences are two maximum length sequences generated by the polynomials $g_1(x) = x^4 + x + 1$ and $g_2(x) = x^4 + x^3 + 1$, [39]. The propagation delay, τ_1 , was estimated by the MUSIC estimator based on both $\{\mathbf{r}(1), \dots, \mathbf{r}(M)\}$ and $\{\mathbf{r}_V(1), \dots, \mathbf{r}_V(M)\}$, where $\mathbf{V}$ is an optimum transform, $M \in \{50, 100, 200, 300, 500\}$, and $L = 4K = 8$. This L represents the largest dimension reduction that can be achieved by any optimum transform.

As seen from Figure 4.3, no performance degradation of the timing estimator was noticed even though the dimension of the received vector was reduced by almost a factor of two.

Recall, however, that the optimum transform requires knowledge of the propaga-

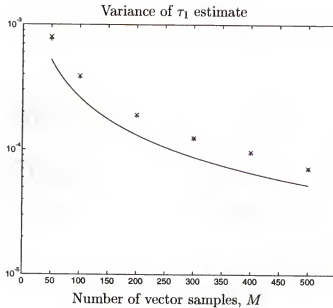


Figure 4.3: CRB: solid line, MUSIC estimate before dimension reduction: (x) and MUSIC estimate after dimension reduction: (+).

tion delays. In other words, we cannot do optimum dimension reduction before we know what we want to estimate. Nevertheless, we can still envision a recursive scheme in which we estimate the propagation delays, compute the optimum transform based on the estimates, and then update the estimates based the new transform. This might still give us a lower total complexity, since many estimators typically require on the order of $(QN)^3$ operations which can be lowered to on the order of L^3 operations per iteration.

4.4 Conclusions

This chapter considered the parameter estimation problem for DS-CDMA systems operating over AWGN channels.

The Cramér-Rao bound (CRB) was presented in Section 4.1. The CRB is of great importance since it is a bound on the best accuracy with which parameters can be estimated by *any* unbiased estimator. In other words, an unbiased estimator

whose variance attains the CRB is optimum. The asymptotic CRB (i.e., the CRB as the observation interval goes to infinity) was also derived. By studying the form of asymptotic CRB, we concluded that there may indeed exist near-far resistant estimators.

In Section 4.2.1, the optimum maximum-likelihood (ML) estimator was defined. Unfortunately, the ML estimator is very complex and thus of limited practical value. To overcome this deficiency, two suboptimum estimators were proposed, the approximate maximum-likelihood (AML) algorithm and the MUSIC algorithm in Section 4.2.2 and 4.2.3, respectively. Of these schemes, the MUSIC algorithm has the lower complexity. In Section 4.2.5, the performance of these estimators was compared with the CRB and the performance of the correlator algorithm. The correlator algorithm, which is analogous to the standard sliding correlator, was shown to work satisfactory when the multiuser interference is low. However, in a near-far situation, the performance is severely degraded. The AML and MUSIC algorithms were shown to have similar performance and to be robust against multiuser interference. Furthermore, they impose no constraints on the transmitted data sequences—in contrast to the sliding correlator—and can therefore be used for tracking as well as for acquisition.

Finally, in Section 4.3, complexity reduction through dimension reduction was proposed. The optimum transform, i.e., the transform that conserves the CRB, was derived. The optimum transform requires knowledge about all users' code waveforms and propagation delays. Thus, since the propagation delays are unknown, the optimum transform is unknown. However, as mentioned in Section 4.3.2, we can still use the optimum transform in a practical manner.

CHAPTER 5

SENSITIVITY ANALYSIS OF THE DECORRELATING RECEIVER

In this chapter we will present a sensitivity analysis and simulation of the popular decorrelator detector with respect to propagation delay estimation errors. The notation used in this chapter is different from the notation in the rest of this dissertation. The reason for this is to be consistent with the notation that is widely used in the literature. This chapter is self-contained, and author hopes that the notational change will not be too confusing.

The original work in this chapter is (1) the extension of the Lupas-Verdú system model to account for propagation delay estimation errors, (2) the proof of the existence of a near-far resistant detector in the presence of imperfect propagation delay estimation, and (3) the experimental results reported in Section 5.5.

5.1 Introduction

The *decorrelating receiver* of Lupas and Verdú [24, 22] as well as several *interference cancellation* schemes [11] have gained considerable interest lately. The reasons for this are that the standard matched filter receiver will not function satisfactory in a near-far environment, and that the complexity of the optimum receiver renders it unfeasible [57].

Near-far resistant structures are typically dependent on accurate side-information, e.g., the number of users, each user's code waveform, received power, carrier phase, and propagation delay. The number of users and their code waveforms might be readily available, but the other parameters must be estimated. It is therefore of interest

to evaluate the performance of the receivers when the available side-information is contaminated with estimation errors.

In this chapter we will study the decorrelating receiver. In its original form, the decorrelator requires knowledge of the number of users, each user's code waveform, received carrier phase, and propagation delay. For reasons explained below, we will assume that the phase is estimated perfectly. We will, however, allow the propagation delay estimates to contain estimation errors, and derive expressions for the error probability, asymptotic efficiency, and near-far resistance.

5.2 System Model

The objective with this section is to extend the system model of Lupas and Verdú to account for propagation delay estimation errors [22, 57].

The received signal is

$$r(t) = S(t, \mathbf{b}) + n(t) \quad (5.1)$$

where $n(t)$ is a white Gaussian waveform with power spectral density $N_0/2$. In general, a subscript k implies that the quantity is due to the k th user. For example, $b_k(i) \in \{-1, 1\}$ is the k th user's i th bit. The signal part, $S(t, \mathbf{b})$, is found as

$$S(t, \mathbf{b}) = \sum_{i=-M}^M \sum_{k=1}^K b_k(i) \sqrt{w_k(i)} s_k(t - iT - \tau_k) \quad (5.2)$$

where K is the number of users, $N = 2M + 1$ is the number of transmitted bits per user, $w_k(i)$ is the received bit energy, $\tau_k \in [0, T)$ is the propagation delay, and $s_k(t) = 0$ for $t \notin [0, T)$, where T is the bit duration. Furthermore, the code waveform, $s_k(t)$, is normalized,

$$\int_{-\infty}^{\infty} s_k^2(t) dt = 1 \quad (5.3)$$

The vector of transmitted bits, $\mathbf{b} \in \{-1, 1\}^{NK}$, is defined as

$$\mathbf{b} = [\mathbf{b}^T(-M) \quad \cdots \quad \mathbf{b}^T(M)]^T \in \{-1, 1\}^{NK} \quad (5.4)$$

$$\mathbf{b}(i) = [b_1(i) \quad \cdots \quad b_K(i)]^T \in \{-1, 1\}^K \quad (5.5)$$

The receiver front-end consists of a bank of matched filters. The output of the k th user's matched filter sampled at the i th bit interval is

$$y_k(i) = \int_{iT+\tau_k}^{(i+1)T+\tau_k} r(t) s_k(t - iT - \tau_k) dt \quad (5.6)$$

It is easy to show that $\{y_k(i)\}$ for $k = 1, \dots, K$ and $i = -M, \dots, M$ is a sufficient statistic for detecting $\mathbf{b}$.

Note that we need to form $s_k(t - \tau_k)$ in order to perform matched filtering. This means that we need to know the carrier phase [which is implicitly included in $s_k(t)$] and τ_k . In practice, the phase and propagation delay are unknown and must be estimated. We choose to ignore the phase error since we are mostly concerned with the near-far problem. This is justified by the following argument. Suppose that all users' received carrier phases are equal. Consider the signal after down-mixing, i.e., after have multiplied the received signal with a replica of the carrier (offset by the estimated phase) and retained the baseband signal (the lowpass portion of the resulting signal). The down-mixing procedure only scales the users' contributions to the received signal by a common factor. The relative power levels between the users are therefore unaffected by the phase estimation error, and the interference situation is unchanged. However, the signal-to-noise ratio will be affected, but since we assume that the system is interference limited, we ignore this problem.

Let the estimate of τ_k be denoted by $\hat{\tau}_k$. We assume that the estimation errors are such that $\hat{\tau}_k \in [0, T)$. If we assume a correlator-style matched filter, the sampled matched filter output is

$$\begin{aligned} y_k(i) &= \int_{iT+\hat{\tau}_k}^{(i+1)T+\hat{\tau}_k} r(t) s_k(t - iT - \hat{\tau}_k) dt \\ &= n_k(i) + \int_{-\infty}^{\infty} S(t, \mathbf{b}) s_k(t - iT - \hat{\tau}_k) dt \end{aligned} \quad (5.7)$$

where

$$n_k(i) = \int_{-\infty}^{\infty} n(t) s_k(t - iT - \hat{\tau}_k) dt \quad (5.8)$$

is a zero-mean Gaussian random variable with crosscorrelation

$$\mathbb{E}[n_k(i)n_l(j)] = \frac{N_0}{2} \int_{-\infty}^{\infty} s_k(t - iT - \hat{\tau}_k) s_l(t - jT - \hat{\tau}_l) dt \quad (5.9)$$

If we define

$$\hat{R}_{kj}(l) = \int_{-\infty}^{\infty} s_k(t - \hat{\tau}_k) s_j(t + lT - \hat{\tau}_j) dt \quad (5.10)$$

$$\hat{R}_{kj}(l) = \int_{-\infty}^{\infty} s_k(t - \hat{\tau}_k) s_j(t + lT - \tau_j) dt \quad (5.11)$$

then $\mathbb{E}[n_k(i)n_l(j)] = \hat{R}_{kl}(i-j)N_0/2$. Recall that $s_k(t) = 0$ for $t \notin [0, T)$, which implies that $\hat{R}_{kj}(l) = \hat{R}_{kj}(l) = 0$ for $|l| > 1$. It is easy to show that

$$\begin{aligned} y_k(i) &= n_k(i) + \sum_{j=-M}^M \sum_{l=1}^K b_l(j) \sqrt{w_l(j)} \hat{R}_{kl}(i-j) \\ &= n_k(i) + \sum_{l=1}^K b_l(i-1) \sqrt{w_l(i-1)} \hat{R}_{kl}(1) + \sum_{l=1}^K b_l(i) \sqrt{w_l(i)} \hat{R}_{kl}(0) \\ &\quad + \sum_{l=1}^K b_l(i+1) \sqrt{w_l(i+1)} \hat{R}_{kl}(-1) \end{aligned} \quad (5.12)$$

where we assume that $b_l(-(M+1)) = b_l(M+1) = 0$. Following the same steps as Lupas [22], we can write

$$\mathbf{y} = \hat{\mathcal{R}} \mathbf{W} \mathbf{b} + \mathbf{n} \quad (5.13)$$

where we have defined

$$\mathbf{y} = [\mathbf{y}^T(-M) \quad \cdots \quad \mathbf{y}^T(M)]^T \in \mathbb{R}^{NK} \quad (5.14)$$

$$\mathbf{y}(i) = [y_1(i) \quad \cdots \quad y_K(i)]^T \in \mathbb{R}^K \quad (5.15)$$

$$\mathbf{n} = [\mathbf{n}^T(-M) \quad \cdots \quad \mathbf{n}^T(M)]^T \in \mathbb{R}^{NK} \quad (5.16)$$

$$\mathbf{n}(i) = [n_1(i) \quad \cdots \quad n_K(i)]^T \in \mathbb{R}^K \quad (5.17)$$

$$\mathbf{W} = \text{diag}(\mathbf{W}(-M), \dots, \mathbf{W}(M)) \in \mathbb{R}^{NK \times NK} \quad (5.18)$$

$$\mathbf{W}(i) = \text{diag}(\sqrt{w_1(i)}, \dots, \sqrt{w_K(i)}) \in \mathbb{R}^{K \times K} \quad (5.19)$$

$$\hat{\mathcal{R}} = \begin{bmatrix} \hat{\mathbf{R}}(0) & \hat{\mathbf{R}}(-1) & 0 & \cdots & 0 \\ \hat{\mathbf{R}}(1) & \hat{\mathbf{R}}(0) & \hat{\mathbf{R}}(-1) & & \vdots \\ 0 & \hat{\mathbf{R}}(1) & \hat{\mathbf{R}}(0) & \ddots & 0 \\ \vdots & & \ddots & \ddots & \hat{\mathbf{R}}(-1) \\ 0 & \cdots & 0 & \hat{\mathbf{R}}(1) & \hat{\mathbf{R}}(0) \end{bmatrix} \quad (5.20)$$

where the (i, j) th element of the matrices $\hat{\mathbf{R}}(l) \in \mathbb{R}^{K \times K}$ and $\hat{\hat{\mathbf{R}}}(l) \in \mathbb{R}^{K \times K}$ are $(\hat{\mathbf{R}}(l))(i, j) = \hat{R}_{ij}(l)$ and $(\hat{\hat{\mathbf{R}}}(l))(i, j) = \hat{\hat{R}}_{ij}(l)$, respectively. The noise vector $\mathbf{n}$ is a Gaussian zero-mean vector with correlation matrix

$$\mathbb{E}[\mathbf{n}\mathbf{n}^T] = \frac{N_0}{2} \hat{\mathcal{R}} \quad (5.21)$$

where matrix $\hat{\hat{\mathcal{R}}}$ is defined as $\hat{\mathcal{R}}$ in (5.20) with the obvious substitutions $\hat{\mathbf{R}}(l) \leftarrow \hat{\hat{\mathbf{R}}}(l)$.

5.3 Analysis of Linear Receivers

A linear receiver is defined in terms of its receiver vector $\mathbf{v} \in \mathbb{R}^{NK}$ and forms the decision of the k th user's i th bit as¹

$$\hat{b}_k(i) = \text{sgn}[\mathbf{v}^T \mathbf{y}] = \text{sgn}[\mathbf{v}^T \hat{\mathcal{R}} \mathbf{W} \mathbf{b} + \mathbf{v}^T \mathbf{n}] \quad (5.22)$$

The noise term $\mathbf{v}^T \mathbf{n}$ is seen to be a zero-mean Gaussian random variable with variance

$$\mathbb{E}[(\mathbf{v}^T \mathbf{n})^2] = \mathbf{v}^T \mathbb{E}[\mathbf{n}\mathbf{n}^T] \mathbf{v} = \frac{N_0}{2} \mathbf{v}^T \hat{\mathcal{R}} \mathbf{v} \quad (5.23)$$

Suppose we let $b_k(i) = 1$, then the probability of error conditioned on $\mathbf{b}$ is

$$P_{e|\mathbf{b}} = Q \left(\frac{\mathbf{v}^T \hat{\mathcal{R}} \mathbf{W} \mathbf{b}}{\sqrt{\frac{N_0}{2} \mathbf{v}^T \hat{\mathcal{R}} \mathbf{v}}} \right) \quad (5.24)$$

and, assuming equally likely independent bits, the probability of error is

$$P_e = \sum_{\substack{\mathbf{b} \in \{-1, 1\}^{NK} \\ b_k(i)=1}} P_{e|\mathbf{b}} \Pr[\mathbf{b} \mid b_k(i) = 1]$$

¹The receiver vector is, of course, allowed to change with k and i . However, the dependence of $\mathbf{v}$ on k and i is, for notational simplicity, not explicitly indicated.

$$= \sum_{\substack{\mathbf{b} \in \{-1,1\}^{NK} \\ b_k(i)=1}} \frac{1}{2^{NK-1}} Q \left(\frac{\mathbf{x}^T \mathbf{W} \mathbf{b}}{\sqrt{\frac{N_0}{2} \mathbf{v}^T \hat{\mathcal{R}} \mathbf{v}}} \right) \quad (5.25)$$

$$\mathbf{x}^T = \mathbf{v}^T \hat{\mathcal{R}} = [\mathbf{x}^T(-M) \quad \cdots \quad \mathbf{x}^T(M)]^T \quad (5.26)$$

$$\mathbf{x}(i) = [x_1(i) \quad \cdots \quad x_K(i)]^T \quad (5.27)$$

To compute the probability of error we need evaluate and sum 2^{NK-1} Q -functions. This clearly gets out of hand even for rather small N and K . However, if we are mostly concerned with the near-far problem, we can also measure performance in terms of the asymptotic efficiency and the near-far resistance [24]. The asymptotic efficiency for the k th user's i th bit is defined as

$$\eta_{k,i} = \sup \left\{ 0 \leq r \leq 1 : \lim_{N_0 \rightarrow 0} P_e / Q \left(\sqrt{\frac{r 2w_k(i)}{N_0}} \right) < \infty \right\} \quad (5.28)$$

In general, $\eta_{k,i}$ will be dependent on $w_l(j)$, since P_e is a function of $w_l(j)$. We therefore define the near-far resistance of a receiver, $\bar{\eta}_{k,i}$, as the worst case asymptotic efficiency

$$\bar{\eta}_{k,i} = \inf_{\substack{w_l(j) \geq 0 \\ w_k(i) \text{ const.}}} \eta_{k,i} \quad (5.29)$$

The inner product $\mathbf{x}^T \mathbf{W} \mathbf{b}$ can be written as

$$\mathbf{x}^T \mathbf{W} \mathbf{b} = x_k(i) \sqrt{w_k(i)} + \sum_{\substack{j \in \{-M, \dots, M\} \\ l \in \{1, \dots, K\} \\ (l,j) \neq (k,i)}} x_l(j) \sqrt{w_l(j)} b_l(j) \quad (5.30)$$

As $N_0 \rightarrow 0$, the error probability will be dominated by the Q -function with the smallest argument. The smallest argument squared is

$$\begin{aligned} & \frac{2}{N_0 \mathbf{v}^T \hat{\mathcal{R}} \mathbf{v}} \left(x_k(i) \sqrt{w_k(i)} - \sum_{\substack{j \in \{-M, \dots, M\} \\ l \in \{1, \dots, K\} \\ (l,j) \neq (k,i)}} |x_l(j)| \sqrt{w_l(j)} \right)^2 \\ &= \frac{2w_k(i)}{N_0} \frac{1}{\mathbf{v}^T \hat{\mathcal{R}} \mathbf{v}} \left(x_k(i) - \sum_{\substack{j \in \{-M, \dots, M\} \\ l \in \{1, \dots, K\} \\ (l,j) \neq (k,i)}} |x_l(j)| \sqrt{\frac{w_l(j)}{w_k(i)}} \right)^2 \end{aligned} \quad (5.31)$$

Thus, the asymptotic efficiency is

$$\eta_{k,i} = \frac{1}{\mathbf{v}^T \hat{\mathcal{R}} \mathbf{v}} \max^2 \left\{ 0, x_k(i) - \sum_{\substack{j \in \{-M, \dots, M\} \\ l \in \{1, \dots, K\} \\ (l,j) \neq (k,i)}} |x_l(j)| \sqrt{\frac{w_l(j)}{w_k(i)}} \right\} \quad (5.32)$$

and, clearly, the near-far resistance is

$$\bar{\eta}_{k,i} = \begin{cases} \frac{x_k^2(i)}{\mathbf{v}^T \hat{\mathcal{R}} \mathbf{v}} & \text{if } x_l(j) = 0, \text{ for all } (l, j) \neq (k, i) \\ 0 & \text{otherwise} \end{cases} \quad (5.33)$$

5.4 Analysis of the Decorrelating Receiver

The decorrelating receiver for the case of imperfect matched filtering is defined as the receiver whose receiver vector satisfies $\hat{\mathcal{R}}^T \hat{\mathbf{v}} = \mathbf{u}_{k,i}$, where $\mathbf{u}_{k,i} \in \mathbb{R}^{NK}$ is the unit vector $\mathbf{u}_{k,i}(n) = \delta_K(n - [(i + M)K + k])$. If we assume that $\hat{\mathcal{R}}$ is invertible (which is reasonable for small estimation errors), then $\hat{\mathbf{v}}^T = \mathbf{u}_{k,i}^T \hat{\mathcal{R}}^{-1}$, and $\mathbf{x}^T = \mathbf{v}^T \hat{\mathcal{R}} = \mathbf{u}_{k,i}^T$. The asymptotic efficiency for the decorrelating receiver, using the notation $[\hat{\mathcal{R}}^{-1}]^T = \hat{\mathcal{R}}^{-T}$, is

$$\eta_{k,i} = \frac{1}{\hat{\mathbf{v}}^T \hat{\mathcal{R}} \hat{\mathbf{v}}} = \frac{1}{\mathbf{u}_{k,i}^T \hat{\mathcal{R}}^{-1} \hat{\mathcal{R}} \hat{\mathcal{R}}^{-T} \mathbf{u}_{k,i}} \quad (5.34)$$

Since the asymptotic efficiency is independent of $w_l(j)$ for $(l, j) \neq (k, i)$, the asymptotic efficiency is equivalent to the near-far resistance. This means that it is theoretically possible to achieve near-far resistance even in the presence of imperfect matched filtering. However, note that $\hat{\mathcal{R}}$ is unknown since the propagation delays are unknown. In other words, we do not have the necessary information to form $\hat{\mathbf{v}}$. One possible approximation to $\hat{\mathbf{v}}$ would be to use $\hat{\mathbf{v}} = \hat{\mathcal{R}}^{-1} \mathbf{u}_{k,i}$.

5.5 Numerical Results

In the above it was shown that the decorrelating receiver based on $\hat{\mathbf{v}}$ is not near-far resistant. It is therefore of interest to evaluate just how sensitive this receiver is to

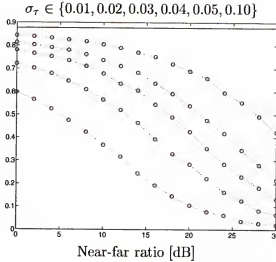


Figure 5.1: Average asymptotic efficiency for decorrelator using $\hat{\mathbf{v}}$: (o) and ideal decorrelator: solid line.

imperfect propagation delay estimates. In this section we will evaluate the probability of error and asymptotic efficiency of the decorrelating receiver for different near-far scenarios and estimation error variances. Because of the complicated nature of (5.25) and (5.32), we resort to Monte-Carlo simulations.

The simulated system was a $K = 3$ user system using Gold code sequences with 15 chips per bit. Each user transmitted $N = 5$ bits. The propagation delays were (randomly) chosen to be $\tau_1 = 2.6253$, $\tau_2 = 8.7261$, and $\tau_3 = 4.5129$, where the chip duration was $T_c = 1$. The propagation delay estimates were assumed to be unbiased and to be independent Gaussian random variables with the same standard deviation σ_τ , i.e., $E[\hat{\tau}_k] = \tau_k$ and $E[(\hat{\tau}_k - \tau_k)^2] = \sigma_\tau^2$. The unbiased assumption is valid if the estimator is allowed to observe the signal for longer than, say, 50 bit intervals (see Section 4.2.5). The assumption of equal standard deviation, however, is not valid if $w_k(i) \neq w_l(j)$. Nevertheless, this assumption was made to facilitate the simulations.

The average asymptotic efficiency of the first user's third bit, $\eta_{1,3}$, is plotted in Figure 5.1. The near-far ratio is defined as $w_l(j)/w_1(1)$ for $l \neq 1$, and $w_k(i)$ was held constant during the data transmission. The asymptotic efficiency was averaged

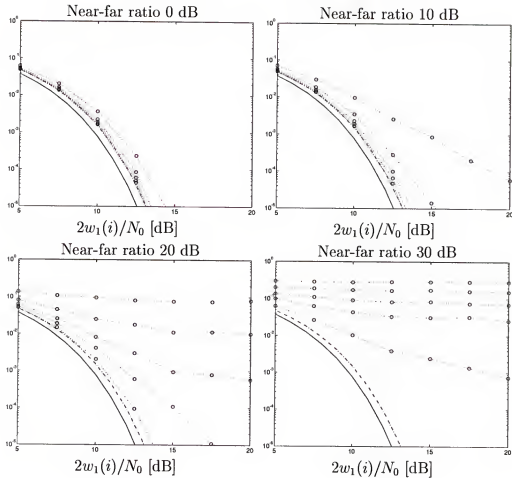


Figure 5.2: Average error probability for the decorrelator using $\hat{\hat{\mathbf{v}}}$ for $\sigma_\tau \in \{0.01, 0.02, 0.03, 0.04, 0.05, 0.1\}$: (o), single-user: solid line, and ideal decorrelator: dash-dotted line.

over 1,500 independent realizations of the propagation delay estimates. The solid line represents $\eta_{1,3}$ for the ideal decorrelator when $\hat{\tau}_k = \tau_k$. Recall that the ideal decorrelator has an asymptotic efficiency that is independent of the near-far ratio. The asymptotic efficiency for the decorrelator based on $\hat{\hat{\mathbf{v}}}$ decreased with the near-far ratio as predicted by (5.32). Furthermore, it is seen that the larger σ_τ was, the worse the asymptotic efficiency became.

Similar conclusions regarding the average error probability are drawn from the plots presented in Figure 5.2. The probability of error for the first user's third bit was

averaged over 500 independent realizations of the propagation delay estimates. The solid line represents the single-user system, i.e., $Q(\sqrt{2w_1(i)/N_0})$, and the dash-dotted line is the error probability of the ideal decorrelator. The performance penalty is seen to have increased with the near-far ratio and with σ_τ . As predicted from (5.25) and (5.33), the error probability of the decorrelator exhibited an error floor, i.e., an irreducible error as the SNR goes to infinity. The position of the error floor is determined by the near-far ratio and σ_τ .

5.6 Conclusions

We have demonstrated that the decorrelating receiver will not be near-far resistant in the presence of imperfect propagation delay estimates. The performance will be impaired by both intersymbol interference and multiple access interference. To maintain acceptable performance, we must keep the variance of the estimation errors small. This constraint becomes more strict as the power of the interfering users become higher. Conventional propagation delay estimation does not work too well in a near-far situation, and we might therefore need to use more sophisticated estimation algorithms.

CHAPTER 6

SYSTEM MODEL FOR FADING CHANNELS

In this chapter we will present a model for a fading channel and derive the equivalent discrete-time complex vector model.

6.1 Continuous-Time Channel Model

We will adopt the channel model proposed by Turin [54]. The model is intuitively pleasing and has been shown to provide a good fit to measured channels [53, 14]. The channel for the k th user is modeled as a time-varying filter with impulse response $h_k(\tau, t)$,

$$h_k(\tau, t) = \sum_{r=0}^{R_k(t)} \alpha_{k,r}(t) \delta(\tau - \tau_{k,r}(t)) \quad (6.1)$$

where $\delta(t)$ is the *Dirac delta function*, $R_k(t)$ is the *number of paths* (or *rays*), $\alpha_{k,r}(t)$ is the *path gain* (or *fading process*), and $\tau_{k,r}(t)$ is the *propagation delay*. We assume that the channel is slowly varying compared with the time for which we are observing the received signal. In particular, we assume that the number of paths and the propagation delays are fixed, i.e., $R_k(t) = R_k$ and $\tau_{k,r}(t) = \tau_{k,r}$.

The channel of the k th user is depicted in Figure 6.1. The r th ray is the transmitted signal delayed with $\tau_{k,r}$ and multiplied with the fading process $\alpha_{k,r}(t)$. The contribution to the received signal from user k can be written as

$$r_k(t) = \sum_{r=1}^{R_k} \alpha_{r,k}(t) s_k(t - \tau_{k,r}) \sqrt{2P_k} \cos(\omega_c t + \theta_{k,r}) \quad (6.2)$$

where $\theta_{k,r} = \theta'_k - \omega_c \tau_{k,r}$, and where $s_k(t)$ is defined by (2.1). We assume that the delays are such that $\tau_{k,r} \in [0, T)$ for $k = 1, 2, \dots, K$ and $r = 1, 2, \dots, R_k$, and that

the propagation delays for a particular user's rays are spaced with more than T_i . Furthermore, we will make the following assumptions:

- $\{\alpha_{k,r}(t)\}$, for $k = 1, 2, \dots, K$ and $r = 1, 2, \dots, R_k$, are independent wide-sense stationary random processes.
- The fading processes vary slowly in time compared with the symbol duration. That is, $\alpha_{r,k}(t) \approx \alpha_{r,k}(mT)$, for $t \in [mT, (m+1)T)$.
- The noise, carrier phase, data, and fading processes are statistically independent.

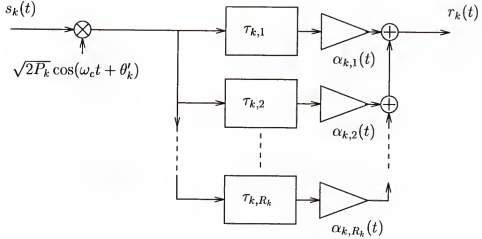


Figure 6.1: Channel model for the k th user

6.2 Discrete-Time Complex Vector System Model

The received signal can be written as

$$r(t) = n(t) + \sum_{k=1}^K \sum_{r=1}^{R_k} \alpha_{k,r}(t) s_k(t - \tau_{k,r}) \sqrt{2P_k} \cos(\omega_c t + \theta_{k,r}) \quad (6.3)$$

where $n(t)$ is a white Gaussian noise waveform with two-sided power spectral density $N_0/2$.

The receiver front-end is the same as in the AWGN channel case, i.e., a standard IQ-stage followed by an integrate-and-dump section, as depicted in Figure 2.1. Ignoring double frequency terms, the equivalent complex received sequence, $r(l) = r_I(l) + jr_Q(l)$, may be written as

$$r(l) = n(l) + \sum_{k=1}^K \sum_{r=1}^{R_k} \beta_{k,r}(l \bmod QN) \frac{1}{T_i} \int_{(l-1)T_i}^{lT_i} s_k(t - \tau_{k,r}) dt \quad (6.4)$$

where

$$\beta_{k,r}(m) = \alpha_{k,r}(mT) \sqrt{P_k} \exp(j\theta_{k,r}) \quad (6.5)$$

and where $n(l)$ is a white Gaussian sequence with variance $\sigma^2 = \mathbb{E}[|n(l)|^2] = N_0/T_i$. The received vector, $\mathbf{r}(m) \in \mathbb{C}^{QN}$, and the noise vector, $\mathbf{n}(m) \in \mathbb{C}^{QN}$, are defined by (2.5) and (2.6), respectively. The noise vector will have the same distribution as in the AWGN case, i.e., $\mathbf{n}(m)$ is a complex Gaussian random vector with zero mean and second moment matrices

$$\mathbb{E}[\mathbf{n}(p)\mathbf{n}^*(q)] = \sigma^2 \mathbf{I}_{QN} \delta_K(p - q), \quad \mathbb{E}[\mathbf{n}(p)\mathbf{n}^T(q)] = \mathbf{0} \quad (6.6)$$

After some straightforward calculations, we can formulate the contribution from the k th user to $\mathbf{r}(m)$ as $\mathbf{r}_k(m) = \mathbf{A}_k \mathbf{s}_k(m)$, where

$$\mathbf{A}_k = [\mathbf{a}_{k,1} \quad \cdots \quad \mathbf{a}_{k,2R_k}] \in \mathbb{R}^{QN \times 2R_k} \quad (6.7)$$

$$\mathbf{s}_k(m) = [s_{k,1}(m) \quad \cdots \quad s_{k,2R_k}(m)]^T \in \mathbb{C}^{2R_k} \quad (6.8)$$

The elements of $\mathbf{s}_k(m)$ are

$$s_{k,2r-1}(m) = \beta_{k,r}(m) z_{2k-1}(m) \quad (6.9)$$

$$s_{k,2r}(m) = \beta_{k,r}(m) z_{2k}(m) \quad (6.10)$$

where $z_{2k-1}(m)$ and $z_{2k}(m)$ are defined by (2.9) and (2.10).

The columns of $\mathbf{A}_k$ are the contributions from the k th user's R_k rays to $\mathbf{r}(m)$. For instance, if $d_k(m) = d_k(m-1)$, the k th user's r th ray will contribute the vector $\beta_{k,r}(m)d_k(m)\mathbf{a}_{k,2r-1}$. On the other hand, if $d_k(m) = -d_k(m-1)$, the contribution will be $\beta_{k,r}(m)d_k(m)\mathbf{a}_{k,2r}$. The form of $\mathbf{a}_{k,2r}$ and $\mathbf{a}_{k,2r-1}$ are defined by $\tau_{k,r}$ and $b_k(t)$ as

$$\mathbf{a}_{k,2r-1} = \left[\frac{\delta_{k,r}}{T_i} \mathbf{D}(p_{k,r} + 1, 1) + \left(1 - \frac{\delta_{k,r}}{T_i} \right) \mathbf{D}(p_{k,r}, 1) \right] \mathbf{c}_k \quad (6.11)$$

$$\mathbf{a}_{k,2r} = \left[\frac{\delta_{k,r}}{T_i} \mathbf{D}(p_{k,r} + 1, -1) + \left(1 - \frac{\delta_{k,r}}{T_i} \right) \mathbf{D}(p_{k,r}, -1) \right] \mathbf{c}_k \quad (6.12)$$

where $\tau_{k,r} = p_{k,r}T_i + \delta_{k,r}$ is such that $p_{k,r}$ is an integer and $\delta_{k,r} \in [0, T_i)$, and where $\mathbf{c}_k$ and $\mathbf{D}(r, s)$ are defined in (2.14) and (2.16), respectively.

The received vector, $\mathbf{r}(m)$, may be expressed as

$$\mathbf{r}(m) = \mathbf{A}\mathbf{s}(m) + \mathbf{n}(m) \quad (6.13)$$

where

$$\mathbf{A} = [\mathbf{A}_1 \quad \mathbf{A}_2 \quad \cdots \quad \mathbf{A}_K] \in \mathbb{R}^{QN \times 2R} \quad (6.14)$$

$$\mathbf{s}(m) = [\mathbf{s}_1^T(m) \quad \mathbf{s}_2^T(m) \quad \cdots \quad \mathbf{s}_K^T(m)]^T \in \mathbb{C}^{2R} \quad (6.15)$$

and where R is *total number of rays*

$$R = \sum_{k=1}^K R_k \quad (6.16)$$

CHAPTER 7

PARAMETER ESTIMATION FOR FADING CHANNELS

In this chapter, the estimation problem from Chapter 4 is generalized to fading channels. The extension to fading channels would be trivial if the fading processes, $\{\alpha_{k,r}(t)\}$, were constant in time. However, this is seldom the case in practice. There are basically two ways in which we can extend the problem.

The first approach is to assume certain distributions for the fading processes, define a set of parameters that determines these distributions, and augment the parameter vector from the AWGN case with those parameters. The problem with this approach is that we must assume particular distributions for the fading processes. This choice is dependent on the application and channel. Moreover, even if we concentrate on a certain environment and application, the choice of distributions is not obvious [17].

The second approach is consider members of the set

$$\{\beta_{k,r}(m) : m = 1, 2, \dots, M; k = 1, 2, \dots, K; r = 1, 2, \dots, R_k\} \quad (7.1)$$

to be unknown and deterministic parameters. The parameter vector is then formed as the $\boldsymbol{\tau}$ vector augmented with σ^2 and the elements of $\{\beta_{k,r}(m)\}$. The advantage with this approach is that we do not need to choose any particular distribution for the fading processes, and the treatment is thus more general. The disadvantage is that the parameter vector is much larger now. In addition to σ^2 and $\boldsymbol{\tau}$, we have MKR complex parameters in $\{\beta_{k,r}(m)\}$. Note that the number of parameters is proportional to the length of the observation interval. This implies that the size of

the Fisher information matrix will grow with the observation interval. Nevertheless, we will adopt this approach due to its generality.

The original work of this chapter is (1) the formulation of the parameter estimation problem for fading channels, (2) the derivation of the Cramér-Rao bound, (3) the formulation of the maximum-likelihood, approximative maximum-likelihood, and MUSIC estimators, and (4) the derivation of the optimum dimension reduction transform.

7.1 Cramér-Rao Bound

The derivation of the Cramér-Rao bound (CRB) is somewhat lengthy, and we therefore defer the details to Appendix A.

The basic assumptions made for the AWGN case in Section 4.1 are made here as well; σ^2 and $\tau_{k,r}$ are assumed to be unknown and deterministic parameters that are fixed during the observation interval. We assume that $\tau_{k,r} \neq pT_i$ for an integer p , which guarantees that $\mathbf{A}$ is differentiable with respect to $\tau_{k,r}$ (see Section A.1). Furthermore we assume that $\{\beta_{k,r}(m)\}$ are unknown and deterministic for $m = 1, 2, \dots, M$, $k = 1, 2, \dots, K$, and $r = 1, 2, \dots, R_k$.

We denote the real part of a complex quantity x by $\tilde{x}$ and the imaginary part by $\tilde{\tilde{x}}$. The parameter vector $\phi \in \mathbb{R}^{1+2RM+R}$ is defined as

$$\phi = [\sigma^2 \quad \beta^T \quad \tau^T]^T \quad (7.2)$$

where

$$\beta = [\check{\beta}^T(1) \quad \check{\beta}^T(1) \quad \check{\beta}^T(2) \quad \check{\beta}^T(2) \quad \dots \quad \check{\beta}^T(M) \quad \check{\beta}^T(M)]^T \in \mathbb{R}^{2RM} \quad (7.3)$$

$$\beta(m) = [\beta_1^T(m) \quad \beta_2^T(m) \quad \dots \quad \beta_{R_k}^T(m)]^T \in \mathbb{R}^R \quad (7.4)$$

$$\beta_k(m) = [\beta_{k,1}(m) \quad \beta_{k,2}(m) \quad \dots \quad \beta_{k,R_k}(m)]^T \in \mathbb{R}^{R_k} \quad (7.5)$$

and

$$\boldsymbol{\tau} = [\boldsymbol{\tau}_1^T \quad \boldsymbol{\tau}_2^T \quad \cdots \quad \boldsymbol{\tau}_K^T]^T \in \mathbb{R}^R \quad (7.6)$$

$$\boldsymbol{\tau}_k = [\tau_{k,1} \quad \tau_{k,2} \quad \cdots \quad \tau_{k,R_k}]^T \in \mathbb{R}^{R_k} \quad (7.7)$$

The Fisher information matrix, $\mathbf{J} \in \mathbb{R}^{(1+2RM+R) \times (1+2RM+R)}$, is found as

$$\mathbf{J} = \mathbb{E} \left[\left(\frac{\partial \ln L(\mathbf{r})}{\partial \boldsymbol{\phi}} \right) \left(\frac{\partial \ln L(\mathbf{r})}{\partial \boldsymbol{\phi}} \right)^T \right] \quad (7.8)$$

where $\ln L(\mathbf{r})$ is the log-likelihood function (conditioned on $\mathbf{z} = [\mathbf{z}^T(1) \cdots \mathbf{z}^T(M)]^T$) of $\mathbf{r} = [\mathbf{r}^T(1) \cdots \mathbf{r}^T(M)]^T$ with respect to $\boldsymbol{\phi}$. Thus, the resulting CRB should be interpreted as being conditioned on both the data sequences and the fading processes.

It is easy to show that the likelihood function conditioned on $\mathbf{z}$ is

$$L(\mathbf{r}) = \frac{1}{(2\pi\sigma^2)^{MQN}} \exp \left(-\frac{1}{\sigma^2} \sum_{m=1}^M \|\mathbf{r}(m) - \mathbf{A}\mathbf{s}(m)\|^2 \right) \quad (7.9)$$

and the log-likelihood function (ignoring constants independent of $\boldsymbol{\phi}$) is

$$\ln L(\mathbf{r}) = -MQN \ln \sigma^2 - \frac{1}{\sigma^2} \sum_{m=1}^M \|\mathbf{r}(m) - \mathbf{A}\mathbf{s}(m)\|^2 \quad (7.10)$$

In Appendix A it is shown that

$$\frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} = -\frac{MQN}{\sigma^2} + \frac{1}{\sigma^4} \sum_{m=1}^M \|\mathbf{n}(m)\|^2 \quad (7.11)$$

$$\frac{\partial \ln L(\mathbf{r})}{\partial \check{\boldsymbol{\beta}}(m)} = \frac{2}{\sigma^2} \text{Re} \{ \mathbf{Z}^*(m) \mathbf{A}^* \mathbf{n}(m) \} \quad (7.12)$$

$$\frac{\partial \ln L(\mathbf{r})}{\partial \boldsymbol{\beta}(m)} = \frac{2}{\sigma^2} \text{Im} \{ \mathbf{Z}^*(m) \mathbf{A}^* \mathbf{n}(m) \} \quad (7.13)$$

$$\frac{\partial \ln L(\mathbf{r})}{\partial \boldsymbol{\tau}} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \{ \boldsymbol{\Gamma}^*(m) \boldsymbol{\Psi}^* \mathbf{n}(m) \} \quad (7.14)$$

where $\boldsymbol{\Psi} \in \mathbb{R}^{QN \times 2R}$ is defined as

$$\boldsymbol{\Psi} = [\boldsymbol{\Psi}_1 \quad \boldsymbol{\Psi}_2 \quad \cdots \quad \boldsymbol{\Psi}_K] \quad (7.15)$$

$$\Psi_k = [\psi_{k,1} \quad \psi_{k,2} \quad \cdots \quad \psi_{k,2R_k}] \in \mathbb{R}^{Q \times R_k} \quad (7.16)$$

$$\psi_{k,2r-1} = \frac{\partial \mathbf{a}_{k,2r-1}}{\partial \tau_{k,r}} = \frac{1}{T_i} [\mathbf{D}(p_{k,r} + 1, 1) - \mathbf{D}(p_{k,r}, 1)] \mathbf{c}_k \quad (7.17)$$

$$\psi_{k,2r} = \frac{\partial \mathbf{a}_{k,2r}}{\partial \tau_{k,r}} = \frac{1}{T_i} [\mathbf{D}(p_{k,r} + 1, -1) - \mathbf{D}(p_{k,r}, -1)] \mathbf{c}_k \quad (7.18)$$

and where $\mathbf{Z}_k(m) \in \{-1, 0, 1\}^{2R_k \times R_k}$, $\mathbf{Z}(m) \in \{-1, 0, 1\}^{2R \times R}$, $\Gamma_k(m) \in \mathbb{C}^{2R_k \times R_k}$, and $\Gamma(m) \in \mathbb{C}^{2R \times R}$ are defined as

$$\mathbf{Z}(m) = \frac{\partial \mathbf{s}(m)}{\partial \beta(m)} = \text{diag}(\mathbf{Z}_1(m), \dots, \mathbf{Z}_K(m)) \quad (7.19)$$

$$(\mathbf{Z}_k(m))(i, j) = \begin{cases} z_{2k-1}(m) & \text{if } i = 2j - 1 \\ z_{2k}(m) & \text{if } i = 2j \\ 0 & \text{otherwise} \end{cases} \quad (7.20)$$

$$\Gamma(m) = \text{diag}(\Gamma_1(m), \dots, \Gamma_K(m)) \quad (7.21)$$

$$(\Gamma_k(m))(i, j) = \begin{cases} s_{k,i}(m) & \text{if } i \in \{2j - 1, 2j\} \\ 0 & \text{otherwise} \end{cases} \quad (7.22)$$

The Fisher information matrix, $\mathbf{J}$, can now be formed by taking the expected value of all possible outer products of the gradients (7.11)–(7.14). After some manipulation we conclude that

$$\mathbf{J} = \begin{bmatrix} \frac{MQN}{\sigma^4} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{ss} & \mathbf{J}_{s\tau} \\ \mathbf{0} & \mathbf{J}_{s\tau}^T & \mathbf{J}_{\tau\tau} \end{bmatrix} \quad (7.23)$$

where

$$\mathbf{J}_{ss} = \text{diag}(\mathbf{J}_{ss}(1), \dots, \mathbf{J}_{ss}(M)) \quad (7.24)$$

$$\mathbf{J}_{ss}(m) = \frac{2}{\sigma^2} \begin{bmatrix} \text{Re}\{\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(m)\} & -\text{Im}\{\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(m)\} \\ -\text{Im}^T\{\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(m)\} & \text{Re}\{\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(m)\} \end{bmatrix} \quad (7.25)$$

and

$$\mathbf{J}_{s\tau} = [\mathbf{J}_{s\tau}^T(1) \quad \cdots \quad \mathbf{J}_{s\tau}^T(M)]^T \quad (7.26)$$

$$\mathbf{J}_{s\tau}(m) = \frac{2}{\sigma^2} \begin{bmatrix} \text{Re}\{\mathbf{Z}^*(m)\mathbf{A}^*\Psi\Gamma(m)\} \\ \text{Im}\{\mathbf{Z}^*(m)\mathbf{A}^*\Psi\Gamma(m)\} \end{bmatrix} \quad (7.27)$$

$$\mathbf{J}_{\tau\tau} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re}\{\Gamma^*(m)\Psi^*\Psi\Gamma(m)\} \quad (7.28)$$

To obtain the CRB for an estimator of τ , we invert the Fisher information matrix

$$\mathbf{J}^{-1} = \begin{bmatrix} \frac{\sigma^4}{MQN} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \times & \times \\ \mathbf{0} & \times & \text{CRB}(\tau) \end{bmatrix} \quad (7.29)$$

where $\text{CRB}(\tau) \in \mathbb{R}^{R \times R}$ is the lower right-hand $R \times R$ block of $\mathbf{J}^{-1}$. In (7.29), $\times$ denote matrices that of no particular interest to us. It is shown in Section A.2 that

$$\text{CRB}(\tau) = [\mathbf{J}_{\tau\tau} - \mathbf{J}_{s\tau}^T \mathbf{J}_{ss}^{-1} \mathbf{J}_{s\tau}]^{-1} \leq \mathbb{E}[(\hat{\tau} - \tau)(\hat{\tau} - \tau)^T] \quad (7.30)$$

where $\hat{\tau}$ is an unbiased estimator of τ . After some algebraic manipulations we obtain

$$\text{CRB}^{-1}(\tau) = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re}\{\Gamma^*(m)\Psi^*\mathbf{P}_{\mathbf{AZ}(m)}^\perp\Psi\Gamma(m)\} \quad (7.31)$$

where $\mathbf{P}_{\mathbf{AZ}(m)}^\perp$ is the projection matrix onto $[\text{range}(\mathbf{AZ}(m))]^\perp$,

$$\mathbf{P}_{\mathbf{AZ}(m)}^\perp = \mathbf{I}_{QN} - \mathbf{AZ}(m)(\mathbf{Z}^*(m)\mathbf{A}^*\mathbf{AZ}(m))^{-1}\mathbf{Z}^*(m)\mathbf{A}^* \quad (7.32)$$

We will now show that the CRB is independent of the near-far problem, i.e., the CRB for an estimator of the propagation delays for k th user's rays is independent of the received power for the other $K - 1$ users. To show this, consider the square-root of the mean power of the k th user's r th ray,

$$q_{k,r} = \sqrt{\mathbb{E}[|\beta_{k,r}(m)|^2]} \quad (7.33)$$

and let $\mathbf{Q} \in \mathbb{R}^{R \times R}$ be defined as

$$\mathbf{Q} = \text{diag}(\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_K) \quad (7.34)$$

$$\mathbf{Q}_k = \text{diag}(q_{k,1}, q_{k,2}, \dots, q_{k,R_k}) \in \mathbb{R}^{R_k \times R_k} \quad (7.35)$$

Moreover, let $\mathbf{\Gamma}_{\text{norm}}(m)\mathbf{Q} = \mathbf{\Gamma}(m)$. Now, if all rays arrive with unit mean power, the CRB is

$$\text{CRB}_{\text{norm}}^{-1}(\boldsymbol{\tau}) = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re}\{\mathbf{\Gamma}_{\text{norm}}^*(m)\boldsymbol{\Psi}^*\mathbf{P}_{\mathbf{Az}(m)}^\perp\boldsymbol{\Psi}\mathbf{\Gamma}_{\text{norm}}(m)\} \quad (7.36)$$

From (7.31) and (7.36) we see that

$$\text{CRB}(\boldsymbol{\tau}) = \mathbf{Q}^{-1}\text{CRB}_{\text{norm}}(\boldsymbol{\tau})\mathbf{Q}^{-1} \quad (7.37)$$

and from this equation we conclude that the CRB on an estimator of a particular ray's propagation delay—given by the appropriate diagonal element of $\text{CRB}(\boldsymbol{\tau})$ —is independent of the mean power of the other rays. In other words, *the CRB is independent of the near-far problem*. This result is important since it tells us that there may exist near-far resistant estimators of $\boldsymbol{\tau}$.

7.2 Estimation Algorithms

The algorithms presented here are rather straightforward modifications of the estimators proposed in Section 4.2.

7.2.1 Maximum-Likelihood

The maximum-likelihood (ML) estimator for the fading channel case is completely analogous to the ML estimator for the AWGN channel presented in Section 4.2.1. The ML estimator is found by minimizing the negative of the log-likelihood function (7.10) with respect to $\mathbf{z}$, defined by (4.23), and $\boldsymbol{\phi}$,

$$\begin{bmatrix} \hat{\boldsymbol{\phi}}^{\text{ML}} \\ \hat{\mathbf{z}}^{\text{ML}} \end{bmatrix} = \arg \min_{\boldsymbol{\phi}, \mathbf{z}} \sum_{m=1}^M \|\mathbf{r}(m) - \mathbf{A}\mathbf{s}(m)\|^2 \quad (7.38)$$

The optimization of the ML cost function is carried out over $1 + 2MR + R$ continuous parameters in $\boldsymbol{\phi}$, and KM discrete parameters in $\mathbf{z}$. The complexity is therefore enormous, even for rather modest values of the observation interval and number of users. We therefore conclude that the ML estimator is of small practical value. We

also note that the ML estimator cannot reach the CRB as long as the oversampling factor $Q < \infty$. The reason for this is that we cannot make consistent estimates of $\boldsymbol{\beta}$ with a finite Q , regardless of the length of the observation interval [46].

7.2.2 Approximative Maximum-Likelihood

We can significantly reduce the complexity of the ML estimator by omitting the dependency of the log-likelihood function on $\mathbf{z}$. This can be done in the same manner as in Section 4.2.2, by considering $\mathbf{s}(m)$ to be continuous, unknown and deterministic. Note that the structure of $\mathbf{s}(m)$ implied by (6.9) and (6.10) is then not used. We can form the approximative maximum-likelihood estimate of $\boldsymbol{\tau}$ as

$$\begin{aligned}\hat{\boldsymbol{\tau}}^{\text{AML}} &= \arg \min_{\boldsymbol{\tau}} \sum_{m=1}^M \|\mathbf{r}(m) - \mathbf{s}(m)\|^2 \Big|_{\mathbf{s}(m) = \mathbf{A}^\dagger(\boldsymbol{\tau})\mathbf{r}(m)} \\ &= \arg \min_{\boldsymbol{\tau}} \text{trace} \left(\mathbf{P}_{\mathbf{A}(\boldsymbol{\tau})}^\perp \hat{\mathbf{R}}_M \right)\end{aligned}\quad (7.39)$$

where $\mathbf{P}_{\mathbf{A}(\boldsymbol{\tau})}^\perp = \mathbf{I}_{QN} - \mathbf{A}(\boldsymbol{\tau})\mathbf{A}^\dagger(\boldsymbol{\tau})$ is the orthogonal projection matrix onto the subspace $[\text{range}(\mathbf{A}(\boldsymbol{\tau}))]^\perp$, and where $\hat{\mathbf{R}}_M$ is the sample correlation matrix

$$\hat{\mathbf{R}}_M = \frac{1}{M} \sum_{m=1}^M \mathbf{r}(m)\mathbf{r}^*(m) \quad (7.40)$$

7.2.3 MUSIC

The MUSIC cost function for fading channels is identical to the MUSIC cost function for the AWGN channel:

$$J_{k,\text{MU}}(\tau) = \frac{\|\hat{\mathbf{E}}_n^* \mathbf{b}_{k,1}(\tau)\|^2}{\|\mathbf{b}_{k,1}(\tau)\|^2} + \frac{\|\hat{\mathbf{E}}_n^* \mathbf{b}_{k,2}(\tau)\|^2}{\|\mathbf{b}_{k,2}(\tau)\|^2} \quad (7.41)$$

where the $\mathbf{b}_{k,1}(\tau)$ and $\mathbf{b}_{k,2}(\tau)$ are defined as

$$\mathbf{b}_{k,1}(\tau) = \left[\frac{\delta}{T_i} \mathbf{D}(p+1, 1) + \left(1 - \frac{\delta}{T_i} \right) \mathbf{D}(p, 1) \right] \mathbf{c}_k \quad (7.42)$$

$$\mathbf{b}_{k,2}(\tau) = \left[\frac{\delta}{T_i} \mathbf{D}(p+1, -1) + \left(1 - \frac{\delta}{T_i} \right) \mathbf{D}(p, -1) \right] \mathbf{c}_k \quad (7.43)$$

where $\tau = pT_i + \delta$ is such that p is an integer and $\delta = [0, T_i)$. $\hat{\mathbf{E}}_n$ is found from the eigenvalue decomposition of $\hat{\mathbf{R}}_M$,

$$\hat{\mathbf{R}}_M = \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s \hat{\mathbf{E}}_s^* + \hat{\mathbf{E}}_n \hat{\mathbf{\Lambda}}_n \hat{\mathbf{E}}_n^* \quad (7.44)$$

where the columns of $\hat{\mathbf{E}}_n$ are the eigenvectors corresponding to the $QN - 2R$ smallest eigenvalues of $\hat{\mathbf{R}}_M$.

The set $\mathcal{T}$ of candidate estimates of τ_k is formed exactly as in Section 4.2.3, and the elements of the MUSIC estimate $\hat{\tau}_k^{\text{MU}}$ are found as the elements of $\mathcal{T}$ that corresponds to the R_k smallest values of $J_{k,\text{MU}}(\tau)$.

The only significant difference between this algorithm and the MUSIC estimator for AWGN channels is that the noise subspace here has dimension $QN - 2R$, and that we pick R_k estimates from $\mathcal{T}$. Recall that in the AWGN case, the noise subspace has dimension $QN - 2K$, and that we pick only one estimate from $\mathcal{T}$.

7.2.4 Numerical Results

The simulated system was a 2-user system with $Q = 1$, $N = 15$ chips per bit, $\text{SNR}_1 = 2E_{b,1}/N_0 = 15$ dB, and 2 rays per user. The rays from each user were assumed to have the same (unit) mean power. The fading processes are set to be the standard Rayleigh fading processes with a Doppler frequency of $2.60 \times 10^{-3}/T = 83.3$ Hz [19, 12]. This Doppler frequency would result if the system had a carrier frequency of 900 MHz, a data rate of 32 kbit/s, and a transmitter-receiver relative speed of 100 km/h.

For each Monte-Carlo run, independent realizations of $\{d_k(m), \alpha_{k,r}(m), \mathbf{n}(m)\}$ were used. The propagation delays were (randomly) set to

$$\boldsymbol{\tau} = T_c [3.70 \quad 5.20 \quad 4.12 \quad 8.16]^T \quad (7.45)$$

The received vector was observed for $M = \{200, 300, 400, 500\}$ symbols, and estimates of $\tau_{1,1}$ were computed by the MUSIC estimator. The variance of the MUSIC algorithm

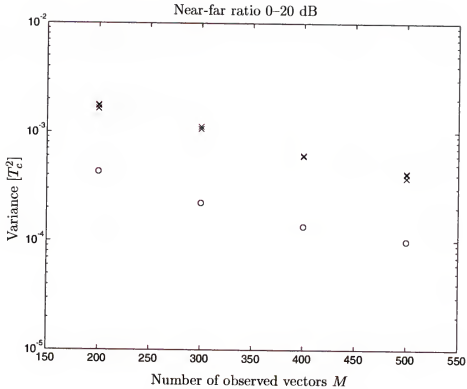


Figure 7.1: Average CRB: (o), variance of the MUSIC estimator: (x).

was estimated as the sample variance found from 10^3 independent Monte-Carlo trials. Since the Cramér-Rao bound is dependent on the realizations of $d_k(m)$ and $\alpha_{k,r}(m)$, the CRB plotted in Figure 7.1 is the average CRB. The near-far ratio, P_2/P_1 , was varied from 0 dB to 20 dB, in order to investigate the near-far resistance of the MUSIC estimator.

The average CRB and the estimated variance of the MUSIC estimator for different near-far ratios are plotted in Figure 7.1. Recall that the CRB is invariant to the near-far ratio. The plot indicates that the MUSIC estimator has the same property.

The strategy outlined in Sections 7.2.3 and 4.2.3 calls for searching all QN bins $\tau \in [pT_i, (p+1)T_i)$ of the cost function. The estimates, $\hat{\tau}_{1,1}^{\text{MU}}$ and $\hat{\tau}_{1,2}^{\text{MU}}$, are picked as the two elements of $\mathcal{T}$ that corresponds to the two smallest values of the cost function.

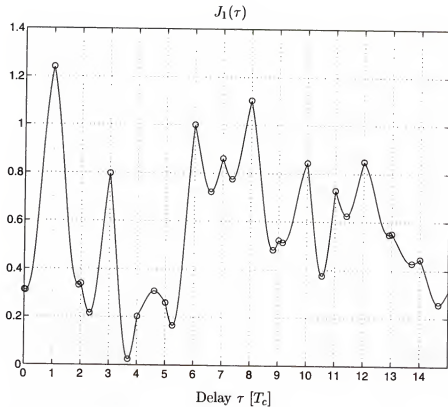


Figure 7.2: MUSIC cost function for $M = 300$ and $P_2/P_1 = 0$ dB.

The procedure is illustrated in Figure 7.2, in which a typical cost function is plotted. The circles indicate the members of $\mathcal{T}$. In this case, the MUSIC algorithm will pick the right bins, i.e., the estimates will be close to $3.7T_c$ and $5.2T_c$. However, if the cost function had been more perturbed, we might have chosen the wrong bin for one or both of $\hat{\tau}_{1,1}^{\text{MU}}$ and $\hat{\tau}_{1,2}^{\text{MU}}$. This type of error, referred to as an *outlier error*, can be quite large. On the other hand, if we choose the correct bin, or restrict the search to be inside the correct bin (as the case will be in tracking), the error is much smaller. To make a meaningful comparison with the CRB, the outlier errors were excluded from the data before computing the variance of $\hat{\tau}_{1,1}^{\text{MU}}$.

The relative number of outlier errors can be found in Table 7.1. Note that these numbers are not very reliable since only 10^3 Monte-Carlo runs were made. However, the table indicates that for large M , there will be only a few outliers. The reason why

outlier errors are frequent for small M is that the ray, whose delay is to be estimated, might be in a deep fade during the entire observation interval, and the signal-to-noise ratio is thereby made very small. Thus, the cost function will be severely perturbed and an outlier error may occur.

Table 7.1: Relative number of outlier errors.

	$M = 200$	300	400	500
$P_2/P_1 = 0$ [dB]	0.245	0.079	0.025	0.005
10 [dB]	0.310	0.090	0.021	0.008
20 [dB]	0.328	0.113	0.030	0.008

7.3 Dimension Reduction

In Section 4.3, we motivated the need for complexity reduction through dimension reduction. Complexity is of course a concern here as well. In this section we will therefore derive the optimum dimension reduction transform for fading channels. As it turns out, the optimum transform is very similar to the optimum transform for the AWGN channel as presented in Section 4.3.

Recall that the optimum transform is defined as a matrix $\mathbf{V} \in \mathbb{C}^{Q^N \times L}$ with the property that the CRB on an estimator of

$$\boldsymbol{\phi}' = [\boldsymbol{\beta}^T \quad \boldsymbol{\tau}^T]^T \quad (7.46)$$

based on the data $\{\mathbf{r}(1), \dots, \mathbf{r}(M)\}$ is the same as the CRB on an estimator based on the data $\{\mathbf{V}^* \mathbf{r}(1), \dots, \mathbf{V}^* \mathbf{r}(M)\}$.

Suppose $\mathbf{V}^* \mathbf{V} = \mathbf{I}_L$, then the reduced dimension received vector may be expressed as

$$\mathbf{r}_V(m) = \mathbf{V}^* \mathbf{r}(m) = \mathbf{A}_V \mathbf{s}(m) + \mathbf{n}_V(m) \quad (7.47)$$

where $\mathbf{A}_V = \mathbf{V}^* \mathbf{A}$ and $\mathbf{n}_V(m) = \mathbf{V}^* \mathbf{n}(m)$ is a complex random Gaussian vector with zero mean and second moment matrices

$$\mathbb{E}[\mathbf{n}_V(m) \mathbf{n}_V^*(k)] = \mathbf{V}^* \mathbb{E}[\mathbf{n}(m) \mathbf{n}^*(k)] \mathbf{V} = \sigma^2 \mathbf{I}_L \delta_K(k - m) \quad (7.48)$$

$$\mathbb{E}[\mathbf{n}_V(m) \mathbf{n}_V^T(k)] = \mathbf{0} \quad (7.49)$$

Thus, the log-likelihood function of $\mathbf{r}_V = [\mathbf{r}_V^T(1) \ \cdots \ \mathbf{r}_V^T(M)]$ with respect to ϕ and conditioned on $\mathbf{z}$ (ignoring constants independent of ϕ) is

$$\ln L(\mathbf{r}_V) = -ML \ln \sigma^2 - \frac{1}{\sigma^2} \sum_{m=1}^M \|\mathbf{r}_V(m) - \mathbf{A}_V \mathbf{s}(m)\|^2 \quad (7.50)$$

We assume that the columns of $\mathbf{A}_V$ are linearly independent for all possible realizations of τ , and we exclude the case when $\delta_{k,r} = 0$ for any $k = 1, 2, \dots, K$ and $r = 1, 2, \dots, R_k$. Then the Fisher information matrix for an unbiased estimator based on $\mathbf{r}_V$ will be as in (7.23), with the appropriate substitutions, $\mathbf{A} \leftarrow \mathbf{A}_V$, $\Psi \leftarrow \Psi_V = \mathbf{V}^* \Psi$, and $QN \leftarrow L$:

$$\mathbf{J}_V = \begin{bmatrix} \frac{ML}{\sigma^4} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{V,ss} & \mathbf{J}_{V,s\tau} \\ \mathbf{0} & \mathbf{J}_{V,s\tau}^T & \mathbf{J}_{V,\tau\tau} \end{bmatrix} \quad (7.51)$$

where

$$\mathbf{J}_{V,ss} = \text{diag}(\mathbf{J}_{V,ss}(1), \dots, \mathbf{J}_{V,ss}(M)) \quad (7.52)$$

$$\mathbf{J}_{V,ss}(m) = \frac{2}{\sigma^2} \begin{bmatrix} \text{Re}\{\mathbf{Z}^*(m) \mathbf{A}_V^* \mathbf{A}_V \mathbf{Z}(m)\} & -\text{Im}\{\mathbf{Z}^*(m) \mathbf{A}_V^* \mathbf{A}_V \mathbf{Z}(m)\} \\ -\text{Im}\{\mathbf{Z}^*(m) \mathbf{A}_V^* \mathbf{A}_V \mathbf{Z}(m)\} & \text{Re}\{\mathbf{Z}^*(m) \mathbf{A}_V^* \mathbf{A}_V \mathbf{Z}(m)\} \end{bmatrix} \quad (7.53)$$

and

$$\mathbf{J}_{V,s\tau} = [\mathbf{J}_{V,s\tau}^T(1) \ \cdots \ \mathbf{J}_{V,s\tau}^T(M)]^T \quad (7.54)$$

$$\mathbf{J}_{V,s\tau}(m) = \frac{2}{\sigma^2} \begin{bmatrix} \text{Re}\{\mathbf{Z}^*(m) \mathbf{A}_V^* \Psi_V \Gamma(m)\} \\ \text{Im}\{\mathbf{Z}^*(m) \mathbf{A}_V^* \Psi_V \Gamma(m)\} \end{bmatrix} \quad (7.55)$$

$$\mathbf{J}_{V,\tau\tau} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re}\{\Gamma^*(m) \Psi_V^* \Psi_V \Gamma(m)\} \quad (7.56)$$

As noted in Theorem 3.5, if $\mathbf{V}$ is such that

$$\mathbf{V}^* \mathbf{V} = \mathbf{I}_L \quad (7.57)$$

$$\text{range}([\mathbf{A} \quad \Psi]) \subseteq \text{range}(\mathbf{V}) \quad (7.58)$$

then $\mathbf{V}\mathbf{V}^* \mathbf{A} = \mathbf{A}$, $\mathbf{V}\mathbf{V}^* \Psi = \Psi$, and

$$\mathbf{A}_V^* \mathbf{A}_V = \mathbf{A}^* \mathbf{V} \mathbf{V}^* \mathbf{A} = \mathbf{A}^* \mathbf{A} \quad (7.59)$$

$$\mathbf{A}_V^* \Psi_V = \mathbf{A}^* \mathbf{V} \mathbf{V}^* \Psi = \mathbf{A}^* \Psi \quad (7.60)$$

$$\Psi_V^* \Psi_V = \Psi^* \mathbf{V} \mathbf{V}^* \Psi = \Psi^* \Psi \quad (7.61)$$

It is now easy to see that if $\mathbf{V}$ satisfy (7.57) and (7.58), then the Fisher information matrices (7.51) and (7.23) are identical (with the exception of the (1,1) element). Thus, the CRB on estimators of ϕ' will be identical in both cases, and $\mathbf{V}$ is optimum.

7.4 Conclusions

In this chapter, we extended the parameter estimation results from Chapter 4 to the fading channels scenario.

The Cramér-Rao bound was presented in Section 7.1, and it was shown that the CRB on an estimator of a particular user's propagation delay is independent of the other users' mean received powers. Thus, the CRB is invariant to the near-far problem, and there may exist near-far robust estimators.

In Section 7.2, the maximum-likelihood (ML), approximative maximum-likelihood (AML), and MUSIC estimators were modified to function in fading channels. The ML estimator will *not* attain the CRB if the oversampling factor is finite. The Cramér-Rao bound for the fading channels might therefore not be as tight as it was for the AWGN channel.

In Section 7.2.4, it was experimentally shown that the MUSIC estimator is robust against the near-far problem. Because of the highly nonlinear nature of the cost

function, the MUSIC algorithm is susceptible to outlier errors. An outlier error occurs when the algorithm picks the wrong bin, i.e., the estimate is chosen to be in $[pT_i, (p+1)T_i)$, where p is an integer, when the true propagation delay lies outside this interval. The resulting error can be very large. However, by allowing the MUSIC estimator to observe the signal for a sufficiently long period of time, the frequency of outlier errors can be made negligible.

The optimum form of dimension reduction is found in Section 7.3. The optimum transform requires knowledge of the $\mathbf{A}$ and Ψ matrices, which, in turn, are defined by the K users' code waveforms and propagation delays for all rays. Thus, since the propagation delays are unknown, the optimum transform is unknown. Nevertheless, we may still be able to use the optimum transform in a practical manner as argued in Section 4.3.2.

CHAPTER 8

SUMMARY AND CONCLUSIONS

8.1 Summary

The near-far problem is a serious threat the performance of asynchronous DS-CDMA communication systems. In a near-far situation, the interference will completely destroy the ability of the standard receiver to detect a weak user. Several near-far resistant receivers exist; however, they are typically rather computationally complex and require accurate synchronization. It is therefore of great importance to find receivers and propagation delay estimators that are both near-far resistant and feasible for implementation.

A convenient vector model for an asynchronous DS-CDMA system operating over additive white Gaussian noise (AWGN) channels was derived in Chapter 2. The system model reveals the underlying structure of the system and is easily implemented in software for computer simulations, e.g., in the MATLAB language [26].

The minimum mean-squared error (MMSE) receiver was defined in Chapter 3. Using the system model from Chapter 2, it was shown that the MMSE receiver is near-far resistant and is optimum when the interference is negligible. The MMSE receiver has relatively low complexity and may therefore be implemented. However, if complexity is still a concern, we can employ dimension reduction to further decrease the amount of hardware and software needed for implementation.

The issue of dimension reduction of the MMSE receiver was investigated in Section 3.6. The optimum transform—in the probability of error sense—was shown to be dependent on side-information which is not assumed to be known to the receiver

(most notably, the propagation delays for all users). Optimum dimension reduction is therefore not directly suitable for implementation. Several suboptimum schemes and one asymptotically optimum scheme were therefore proposed in Section 3.6.

The conclusions drawn in Chapter 3 were that the MMSE receiver is attractive due to its simplicity and near-far properties, and that significant dimension reduction may be obtained if complexity is a greater concern than performance.

To maximize performance we need to synchronize the receiver with the transmitted signals. It was shown in Chapter 5 that the popular near-far resistant decorrelating receiver is very sensitive to synchronization errors. Furthermore, as shown in Chapter 4, the standard algorithm for propagation delay estimation fails in a near-far situation. This indicates the need for accurate near-far robust propagation delay estimation algorithms.

The parameter estimation problem for AWGN channels was treated in Chapter 4. Firstly, we presented the Cramér-Rao bound (CRB). The CRB bounds the accuracy of any unbiased estimator and serves as the optimality criterion. It was shown that the CRB, asymptotically as the observation interval goes to infinity, is independent of the near-far problem. Thus, there is no fundamental reason for why propagation delay estimators should be sensitive to the near-far problem. Secondly, the optimum maximum-likelihood (ML) estimator was derived and shown to be very complex. Two suboptimum algorithms, the approximative maximum-likelihood (AML) and MUSIC estimators, were therefore proposed. These algorithms were experimentally shown to be robust against the near-far problem and to have similar performance. The MUSIC algorithm is less complex and should therefore be the method of choice.

The complexity of the parameter estimators can be reduced by dimension reduction. The optimum linear dimension reduction transform for the parameter estimation problem was derived in Section 4.3. Significant reduction in complexity may be attained by the optimum transform; however, the transform requires knowledge of

the users' (unknown) propagation delays and code waveforms. Thus, the optimum transform is mostly of theoretical value.

The results mentioned above are for a system operating over AWGN channels. However, a radio channel can seldom be accurately modeled as an AWGN channel. In Chapter 6, we therefore extended the system model to fading channels. The fading channel was modeled as a multiray channel with nonuniform spacing between rays. This model is very general and has been used by many other researchers in the field. Based on the fading channel model, the parameter estimation problem was reformulated in Chapter 7. The CRB for fading channels was presented in Section 7.1 and shown to be independent of the near-far problem. It may therefore be possible to find near-far resistant estimators for fading channels.

The ML, AML, and MUSIC estimators were modified to fit the fading channel scenario in Section 7.2. The AML and ML estimators are even more complex in the fading case, and furthermore the ML estimator will not be optimum if the oversampling factor is finite. The MUSIC estimator is only slightly more complex in the fading case and was experimentally shown to be robust against the near-far problem in Section 7.2.4.

Dimension reduction for the parameter estimation problem can be employed for fading channels as well, and the optimum transform was derived in Section 7.3. The optimum transform requires basically the same side-information as in the AWGN case—i.e., the users' code waveforms and propagation delays—plus knowledge of the number of rays in each user's channel.

8.2 Contributions

The main contributions made in this dissertation were:

- Introduction of the discrete-time complex vector model which reveals the underlying structure of a DS-CDMA system and facilitates computer simulations.

- Proofs of the behavior of the MMSE receiver when the noise or interference disappear.
- Derivation of the optimum dimension reduction transform for the MMSE receiver.
- Formulation of several practical dimension reduction transforms for the MMSE receiver that outperform previously known schemes.
- Analysis of the sensitivity of the decorrelating receiver to propagation delay estimation errors.
- Derivation of the Cramér-Rao bound for both AWGN and fading channels.
- Formulation of several practical near-far robust propagation delay estimators for both AWGN and fading channels.
- Derivation of the optimum dimension reduction for the parameter estimation problem for both AWGN and fading channels.

8.3 Future Work

There are many natural extensions to the problems studied in this dissertation.

In Chapter 3, we described the linear-conjugate-linear (LCL) modification to the MMSE receiver. This modification is very attractive, since it gives better performance for a lower complexity. It is therefore of interest to extend the results for the standard MMSE receiver to the LCL-MMSE receiver.

The MMSE receiver was analyzed under the assumption that the “adaptive algorithm” in Figure 3.2 can find the Wiener-Hopf solution exactly. In reality, this will not be exactly true. However, in the AWGN case, we can make the deviation from the Wiener-Hopf solution arbitrarily small by using a *least-mean squared* (LMS) algorithm with a sufficiently small step size [15]. In fading, we ideally would like for

the MMSE receiver to update its receiver vector to fit the interference situation on a vector-sample-by-vector-sample basis. This means that the adaptive filter has to track a time-varying solution. However, for typical channels, the scenario changes too fast for the LMS algorithm. Modification to the signaling format or to the adaptive algorithm may therefore be required. To the author's knowledge, no successful modification has yet been proposed. Attempts have been done by using differential BPSK and the *recursive least-squares* (RLS) algorithm [3, 62]. Differential BPSK is used to decrease the sensitivity of the receiver to the time-varying phase of the desired signal, and the use of the RLS algorithm is proposed in order to improve tracking. However, even though RLS has faster convergence than the LMS algorithm in a stationary environment, it is well-known that RLS has roughly the same tracking ability as LMS in a nonstationary environment [15].

Another complication with the fading channel is that the number of rays is generally unknown. However, in Chapter 7, we assumed that this number was known to the estimation algorithms. In practice, we must estimate the number of rays, e.g., by comparing the singular values of the sample correlation matrix with a threshold. Also, the propagation delays will typically change with time. Naturally we would like to track these changes. Future work will therefore include the study of algorithms for estimating the number of rays as well as a recursive formulation of the MUSIC algorithm for tracking.

The optimum transform for dimension reduction shows the potential for significant complexity reduction. However, since the optimum transform is unknown, it is of interest to formulate practical, but probably suboptimum, dimension reduction transforms and study their performance.

APPENDIX A DERIVATION OF THE CRAMÉR-RAO BOUND

For a general discussion of the *Cramér-Rao bound* (CRB), see Scharf [40] or Van Trees [55]. The derivation of the CRB for the AWGN channel, in Section A.1, and for the fading channels, in Section A.2, follows closely the procedure presented by Stoica and Nehorai [46].

A.1 CRB for the AWGN Channel

We start by repeating some key equations from Chapter 2. The received vector for the m th data interval is

$$\mathbf{r}(m) = \mathbf{A}(\boldsymbol{\tau})\mathbf{B}(\boldsymbol{\gamma}, \boldsymbol{\theta})\mathbf{z}(m) + \mathbf{n}(m) \quad (\text{A.1})$$

where $\mathbf{n}(m)$ is a complex Gaussian random vector with zero mean and second moment matrices

$$\mathbb{E}[\mathbf{n}(p)\mathbf{n}^*(q)] = \sigma^2 \mathbf{I}_{QN} \delta_K(p - q) \quad (\text{A.2})$$

$$\mathbb{E}[\mathbf{n}(p)\mathbf{n}^T(q)] = \mathbf{0} \quad (\text{A.3})$$

Conditioned on $\boldsymbol{\phi}$ and $\mathbf{z}(m)$, the received vector $\mathbf{r}(m)$ will be complex Gaussian vector with mean

$$\mathbb{E}[\mathbf{r}(m) \mid \boldsymbol{\phi}, \mathbf{z}(m)] = \mathbf{A}\mathbf{B}\mathbf{z}(m) \quad (\text{A.4})$$

and second moment matrices

$$\mathbb{E} \left[\left(\mathbf{r}(p) - \mathbf{A}\mathbf{B}\mathbf{z}(p) \right) \left(\mathbf{r}(q) - \mathbf{A}\mathbf{B}\mathbf{z}(q) \right)^* \mid \boldsymbol{\phi}, \mathbf{z}(m) \right] = \sigma^2 \mathbf{I}_{QN} \delta_K(p - q) \quad (\text{A.5})$$

$$\mathbb{E} \left[\left(\mathbf{r}(p) - \mathbf{A}\mathbf{B}\mathbf{z}(p) \right) \left(\mathbf{r}(q) - \mathbf{A}\mathbf{B}\mathbf{z}(q) \right)^T \mid \boldsymbol{\phi}, \mathbf{z}(m) \right] = \mathbf{0} \quad (\text{A.6})$$

The probability density function for $\mathbf{r}(m)$ conditioned on ϕ and $\mathbf{z}(m)$ is called the likelihood function for $\mathbf{r}(m)$ and is denoted by $L(\mathbf{r}(m))$. The likelihood function can be written as

$$\begin{aligned} L(\mathbf{r}(m)) &= \frac{1}{\pi^{QN} \det(\sigma^2 \mathbf{I}_{QN})} \exp \left(-\frac{1}{\sigma^2} \|\mathbf{r}(m) - \mathbf{A}\mathbf{B}\mathbf{z}(m)\|^2 \right) \\ &= \frac{1}{(\pi\sigma^2)^{QN}} \exp \left(-\frac{1}{\sigma^2} \|\mathbf{r}(m) - \mathbf{A}\mathbf{B}\mathbf{z}(m)\|^2 \right) \end{aligned} \quad (\text{A.7})$$

Since $\mathbf{r}(p)$ and $\mathbf{r}(q)$ are uncorrelated complex Gaussian random vectors for $p \neq q$, the likelihood function of $\mathbf{r} = [\mathbf{r}^T(1) \ \mathbf{r}^T(2) \ \dots \ \mathbf{r}^T(M)]^T$ may be written as

$$L(\mathbf{r}) = \prod_{m=1}^M L(\mathbf{r}(m)) \quad (\text{A.8})$$

Taking the natural logarithm of $L(\mathbf{r})$ and ignoring terms that are independent of ϕ yields the log-likelihood function

$$\ln L(\mathbf{r}) = -MQN \ln \sigma^2 - \frac{1}{\sigma^2} \sum_{m=1}^M \|\mathbf{r}(m) - \mathbf{A}\mathbf{B}\mathbf{z}(m)\|^2 \quad (\text{A.9})$$

A.1.1 Differentiating the Log-Likelihood Function

We are now ready to start computing the gradient $\partial \ln L(\mathbf{r}) / \partial \phi$. We will do this in four steps by computing the gradient of $\ln L(\mathbf{r})$ with respect to σ^2 , γ , θ , and τ , respectively, and then form $\partial \ln L(\mathbf{r}) / \partial \phi$ as

$$\frac{\partial \ln L(\mathbf{r})}{\partial \phi} = \begin{bmatrix} \frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} \\ \frac{\partial \ln L(\mathbf{r})}{\partial \gamma} \\ \frac{\partial \ln L(\mathbf{r})}{\partial \theta} \\ \frac{\partial \ln L(\mathbf{r})}{\partial \tau} \end{bmatrix} \quad (\text{A.10})$$

The information matrix (4.4) can be expressed as a 4×4 block matrix where each block is the expected value of the outer product of two gradients. For instance, the (2,3) block is

$$\mathbf{J}_{\gamma\theta} = \mathbb{E} \left[\left(\frac{\partial \ln L}{\partial \gamma} \right) \left(\frac{\partial \ln L}{\partial \theta} \right)^T \right] \quad (\text{A.11})$$

Note that the information matrix is real symmetric (and positive semidefinite), and therefore the (3, 2) block is simply the (2, 3) block transposed.

We will find the following results useful.

RESULT 1 Let $\mathbf{x} \in \mathbb{C}^n$ be a function of $\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \cdots \ \theta_m]^T \in \mathbb{R}^m$, and suppose $\mathbf{A} \in \mathbb{C}^{n \times n}$ is a constant matrix. Then

$$\frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{x}^* \mathbf{A} \mathbf{x} = \left(\frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} \right)^* \mathbf{A} \mathbf{x} + \left(\frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} \right)^T \mathbf{A}^T \bar{\mathbf{x}} \quad (\text{A.12})$$

where

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial \mathbf{x}}{\partial \theta_1} & \frac{\partial \mathbf{x}}{\partial \theta_2} & \cdots & \frac{\partial \mathbf{x}}{\partial \theta_m} \end{bmatrix} \in \mathbb{C}^{n \times m} \quad (\text{A.13})$$

RESULT 2 Let $\mathbf{x} \in \mathbb{C}^n$ be a function of $\boldsymbol{\theta} \in \mathbb{R}^m$, and suppose $\mathbf{a} \in \mathbb{C}^n$ is a constant vector. Then

$$\frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{x}^* \mathbf{a} = \left(\frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} \right)^* \mathbf{a} \quad (\text{A.14})$$

$$\frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{a}^* \mathbf{x} = \left(\frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} \right)^T \bar{\mathbf{a}} \quad (\text{A.15})$$

RESULT 3 Let $\mathbf{X} \in \mathbb{C}^{n \times k}$ be function of $\theta \in \mathbb{R}$, and suppose $\mathbf{a} \in \mathbb{C}^n$ is a constant vector. Then

$$\frac{\partial}{\partial \theta} \mathbf{a}^* \mathbf{X}^* \mathbf{X} \mathbf{a} = \mathbf{a}^* \left(\frac{\partial \mathbf{X}}{\partial \theta} \right)^* \mathbf{X} \mathbf{a} + \mathbf{a}^* \mathbf{X}^* \left(\frac{\partial \mathbf{X}}{\partial \theta} \right) \mathbf{a} \quad (\text{A.16})$$

Computing $\partial \ln L(\mathbf{r}) / \partial \sigma^2$

We start by differentiating the log-likelihood function with respect to σ^2 ,

$$\frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} = -\frac{MQN}{\sigma^2} + \frac{1}{\sigma^4} \sum_{m=1}^M \mathbf{n}^*(m) \mathbf{n}(m) \quad (\text{A.17})$$

Computing $\partial \ln L(\mathbf{r}) / \partial \boldsymbol{\gamma}$

Straightforward differentiation with respect to γ_k yields

$$\frac{\partial \ln L(\mathbf{r})}{\partial \gamma_k} = -\frac{1}{\sigma^2} \sum_{m=1}^M \frac{\partial}{\partial \gamma_k} (\mathbf{r}(m) - \mathbf{A} \mathbf{B} \mathbf{z}(m))^* (\mathbf{r}(m) - \mathbf{A} \mathbf{B} \mathbf{z}(m))$$

$$\begin{aligned}
&= -\frac{1}{\sigma^2} \sum_{m=1}^M \left[\left(-\mathbf{A} \frac{\partial \mathbf{B}}{\partial \gamma_k} \mathbf{z}(m) \right)^* \mathbf{n}(m) + \mathbf{n}^*(m) \left(-\mathbf{A} \frac{\partial \mathbf{B}}{\partial \gamma_k} \mathbf{z}(m) \right) \right] \\
&= \frac{1}{\sigma^2} \sum_{m=1}^M 2 \operatorname{Re} \left\{ \left(\mathbf{A} \frac{\partial \mathbf{B}}{\partial \gamma_k} \mathbf{z}(m) \right)^* \mathbf{n}(m) \right\} \\
&= \frac{1}{\sigma^2} \sum_{m=1}^M 2 \operatorname{Re} \left\{ \mathbf{z}^*(m) \frac{\partial \mathbf{B}^*}{\partial \gamma_k} \mathbf{A}^* \mathbf{n}(m) \right\} \tag{A.18}
\end{aligned}$$

where

$$\frac{\partial \mathbf{B}^*}{\partial \gamma_k}(i, j) = \begin{cases} \exp(-j\theta_k) & \text{if } i = j = 2k - 1 \\ \exp(-j\theta_k) & \text{if } i = j = 2k \\ 0 & \text{otherwise} \end{cases} \tag{A.19}$$

Now,

$$\mathbf{z}^*(m) \frac{\partial \mathbf{B}^*}{\partial \gamma_k} \mathbf{A}^* \mathbf{n}(m) = \exp(-j\theta_k) \begin{bmatrix} z_{2k-1}(m) & z_{2k}(m) \end{bmatrix} \begin{bmatrix} \mathbf{a}_{2k-1}^* \\ \mathbf{a}_{2k}^* \end{bmatrix} \mathbf{n}(m) \tag{A.20}$$

and the gradient with respect to γ can be expressed as

$$\begin{aligned}
\frac{\partial \ln L(\mathbf{r})}{\partial \gamma} &= \frac{1}{\sigma^2} \sum_{m=1}^M 2 \operatorname{Re} \left[\begin{array}{cc} \exp(-j\theta_1) \begin{bmatrix} z_1(m) & z_2(m) \end{bmatrix} \begin{bmatrix} \mathbf{a}_1^* \\ \mathbf{a}_2^* \end{bmatrix} \mathbf{n}(m) \\ \vdots \\ \exp(-j\theta_K) \begin{bmatrix} z_{2K-1}(m) & z_{2K}(m) \end{bmatrix} \begin{bmatrix} \mathbf{a}_{2K-1}^* \\ \mathbf{a}_{2K}^* \end{bmatrix} \mathbf{n}(m) \end{array} \right] \\
&= \frac{1}{\sigma^2} \sum_{m=1}^M 2 \operatorname{Re} \left\{ \mathbf{Z}^*(m) \mathbf{B}_\gamma^* \mathbf{A}^* \mathbf{n}(m) \right\} \tag{A.21}
\end{aligned}$$

where $\mathbf{B}_\gamma \in \mathbb{C}^{2K \times 2K}$ is defined as

$$\mathbf{B}_\gamma = \operatorname{diag}(\exp(j\theta_1), \exp(j\theta_1), \dots, \exp(j\theta_K), \exp(j\theta_K)) \tag{A.22}$$

and $\mathbf{Z}(m) \in \{-1, 0, 1\}^{2K \times K}$ is defined as

$$(\mathbf{Z}(m))(i, k) = \begin{cases} z_{2k-1}(m) & \text{if } i = 2k - 1 \\ z_{2k}(m) & \text{if } i = 2k \\ 0 & \text{otherwise} \end{cases} \tag{A.23}$$

Computing $\partial \ln L(\mathbf{r})/\partial \boldsymbol{\theta}$

The derivation above for the gradient with respect to $\boldsymbol{\gamma}$ works here as well, if we make the substitution $\partial \mathbf{B}/\partial \gamma_k \leftarrow \partial \mathbf{B}/\partial \theta_k$. The form of $\partial \ln L(\mathbf{r})/\partial \boldsymbol{\theta}$ will be similar to the form of $\partial \ln L(\mathbf{r})/\partial \boldsymbol{\gamma}$,

$$\frac{\partial \ln L(\mathbf{r})}{\partial \boldsymbol{\theta}} = \frac{1}{\sigma^2} \sum_{m=1}^M 2 \operatorname{Re} \{ \mathbf{Z}^*(m) \mathbf{B}_{\boldsymbol{\theta}}^* \mathbf{A}^* \mathbf{n}(m) \} \quad (\text{A.24})$$

where $\mathbf{B}_{\boldsymbol{\theta}} \in \mathbb{C}^{2K \times 2K}$ is defined as

$$\mathbf{B}_{\boldsymbol{\theta}} = \operatorname{diag}(j\beta_1, j\beta_1, j\beta_2, j\beta_2, \dots, j\beta_K, j\beta_K) = j\mathbf{B} \quad (\text{A.25})$$

Computing $\partial \ln L(\mathbf{r})/\partial \boldsymbol{\tau}$

Straightforward differentiation yields

$$\begin{aligned} \frac{\partial \ln L(\mathbf{r})}{\partial \tau_k} &= -\frac{1}{\sigma^2} \sum_{m=1}^M -\mathbf{n}(m)^* \frac{\partial \mathbf{A}}{\partial \tau_k} \mathbf{B} \mathbf{z}(m) - \mathbf{z}^*(m) \mathbf{B}^* \frac{\partial \mathbf{A}^*}{\partial \tau_k} \mathbf{n}(m) \\ &= \frac{1}{\sigma^2} \sum_{m=1}^M 2 \operatorname{Re} \left\{ \mathbf{z}^*(m) \mathbf{B}^* \frac{\partial \mathbf{A}^*}{\partial \tau_k} \mathbf{n}(m) \right\} \end{aligned} \quad (\text{A.26})$$

Recall the expression for the columns of $\mathbf{A}$,

$$\mathbf{a}_{2k-1} = \left[\frac{\delta_k}{T_i} \mathbf{D}(p_k+1, 1) + \left(1 - \frac{\delta_k}{T_i} \right) \mathbf{D}(p_k, 1) \right] \mathbf{c}_k \quad (\text{A.27})$$

$$\mathbf{a}_{2k} = \left[\frac{\delta_k}{T_i} \mathbf{D}(p_k+1, -1) + \left(1 - \frac{\delta_k}{T_i} \right) \mathbf{D}(p_k, -1) \right] \mathbf{c}_k \quad (\text{A.28})$$

As seen from these equations the derivative of $\mathbf{A}$ with respect to τ_k , or equivalently with respect to δ_k , might not exist when $\delta_k = 0$. The reason for this is that $\mathbf{a}_{2k}$ resides on a straight line between $\mathbf{D}(p_k+1, -1)\mathbf{c}_k$ and $\mathbf{D}(p_k, -1)\mathbf{c}_k$ when $\tau_k \in [p_k T_i, (p_k+1)T_i)$ and between $\mathbf{D}(p_k+2, -1)\mathbf{c}_k$ and $\mathbf{D}(p_k+1, -1)\mathbf{c}_k$ when $\tau_k \in [(p_k+1)T_i, (p_k+2)T_i)$. Now if $\mathbf{D}(p_k, -1)\mathbf{c}_k$, $\mathbf{D}(p_k+1, -1)\mathbf{c}_k$, and $\mathbf{D}(p_k+2, -1)\mathbf{c}_k$ are not collinear, then the derivative will not exist for $\tau_k = (p_k+1)T_i$. The Cramér-Rao bound is not valid if the derivative does not exist (see, for instance, Van Trees [55]). In order to

proceed, we therefore make the assumption that $\delta_k \neq 0$ for $k = 1, 2, \dots, K$. This is a rather mild restriction, since the assumption only implies that we are not sampling in synchronization with any user. It now follows immediately that

$$\frac{\partial \mathbf{A}}{\partial \tau_k} = \begin{bmatrix} 0 & \cdots & 0 & \frac{\partial \mathbf{a}_{2k-1}}{\partial \tau_k} & \frac{\partial \mathbf{a}_{2k}}{\partial \tau_k} & 0 & \cdots & 0 \end{bmatrix} \quad (\text{A.29})$$

Now if we define

$$\psi_{2k-1} = \frac{\partial \mathbf{a}_{2k-1}}{\partial \tau_k} = \frac{1}{T_i} [\mathbf{D}(p_k + 1, 1) - \mathbf{D}(p_k, 1)] \mathbf{c}_k \quad (\text{A.30})$$

$$\psi_{2k} = \frac{\partial \mathbf{a}_{2k}}{\partial \tau_k} = \frac{1}{T_i} [\mathbf{D}(p_k + 1, -1) - \mathbf{D}(p_k, -1)] \mathbf{c}_k \quad (\text{A.31})$$

then

$$\mathbf{z}^*(m) \mathbf{B}^* \frac{\partial \mathbf{A}^*}{\partial \tau_k} \mathbf{n}(m) = \bar{\beta}_k \begin{bmatrix} z_{2k-1}(m) & z_{2k}(m) \end{bmatrix} \begin{bmatrix} \psi_{2k-1}^* \\ \psi_{2k}^* \end{bmatrix} \mathbf{n}(m) \quad (\text{A.32})$$

and the gradient with respect to $\boldsymbol{\tau}$ can be expressed as

$$\begin{aligned} \frac{\partial \ln L(\mathbf{r})}{\partial \boldsymbol{\tau}} &= \frac{1}{\sigma^2} \sum_{m=1}^M 2 \operatorname{Re} \begin{bmatrix} \bar{\beta}_1 \begin{bmatrix} z_1(m) & z_2(m) \end{bmatrix} \begin{bmatrix} \psi_1^* \\ \psi_2^* \end{bmatrix} \mathbf{n}(m) \\ \vdots \\ \bar{\beta}_K \begin{bmatrix} z_{2K-1}(m) & z_{2K}(m) \end{bmatrix} \begin{bmatrix} \psi_{2K-1}^* \\ \psi_{2K}^* \end{bmatrix} \mathbf{n}(m) \end{bmatrix} \\ &= \frac{1}{\sigma^2} \sum_{m=1}^M 2 \operatorname{Re} \{ \mathbf{Z}^*(m) \mathbf{B}^* \boldsymbol{\Psi}^* \mathbf{n}(m) \} \end{aligned} \quad (\text{A.33})$$

where $\boldsymbol{\Psi} \in \mathbb{R}^{QN \times 2K}$ is defined as

$$\boldsymbol{\Psi} = [\psi_1 \quad \psi_2 \quad \cdots \quad \psi_{2K}] \quad (\text{A.34})$$

A.1.2 Forming the Fisher Information Matrix

We have now computed all the gradients needed to form the Fisher information matrix. To continue, we will need a few results (see [46] for proofs). Suppose $\mathbf{n}(m)$ is defined as in (2.6).

RESULT 4

$$\mathbb{E}[\mathbf{n}^*(m) \mathbf{n}(m) \mathbf{n}^*(k) \mathbf{n}(k)] = \begin{cases} (QN)^2 \sigma^4 & \text{if } m \neq k \\ QN(QN + 1) \sigma^4 & \text{if } m = k \end{cases} \quad (\text{A.35})$$

RESULT 5 For all m, k ,

$$\mathbf{E}[\mathbf{n}^*(m)\mathbf{n}(m)\mathbf{n}^T(k)] = \mathbf{0} \quad (\text{A.36})$$

RESULT 6 Suppose $\mathbf{x}, \mathbf{y} \in \mathbb{C}^{QN}$, then

$$\text{Re}\{\mathbf{x}\} \text{Re}\{\mathbf{y}^T\} = \frac{1}{2} [\text{Re}\{\mathbf{xy}^T\} + \text{Re}\{\mathbf{xy}^*\}] \quad (\text{A.37})$$

$$\text{Im}\{\mathbf{x}\} \text{Im}\{\mathbf{y}^T\} = -\frac{1}{2} [\text{Re}\{\mathbf{xy}^T\} - \text{Re}\{\mathbf{xy}^*\}] \quad (\text{A.38})$$

$$\text{Re}\{\mathbf{x}\} \text{Im}\{\mathbf{y}^T\} = \frac{1}{2} [\text{Im}\{\mathbf{xy}^T\} - \text{Im}\{\mathbf{xy}^*\}] \quad (\text{A.39})$$

The $(1, 1)$ block of the Fisher information matrix is

$$\begin{aligned} \mathbf{E} \left[\left(\frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} \right)^2 \right] &= \mathbf{E} \left[\frac{(MQN)^2}{\sigma^4} - \frac{2MQN}{\sigma^6} \sum_{m=1}^M \mathbf{n}^*(m)\mathbf{n}(m) \right. \\ &\quad \left. + \frac{1}{\sigma^8} \sum_{m=1}^M \sum_{k=1}^M \mathbf{n}^*(m)\mathbf{n}(m)\mathbf{n}^*(k)\mathbf{n}(k) \right] \\ &= \frac{(MQN)^2}{\sigma^4} - \frac{2MQN}{\sigma^6} MQN\sigma^2 \\ &\quad + \frac{1}{\sigma^8} [MQN(QN+1)\sigma^4 + (M^2 - M)(QN)^2\sigma^4] \\ &= \frac{MQN}{\sigma^4} \end{aligned} \quad (\text{A.40})$$

We note, by using Result 5, that $\partial \ln L(\mathbf{r})/\partial \sigma^2$ is uncorrelated with all other gradients. Thus, the $(1, 2)$, $(1, 3)$, and $(1, 4)$ blocks of the information matrix will be zero.

To find the remaining blocks of the information matrix, we need to compute the expected values of all possible outer products of the gradients (A.21), (A.24), and (A.33). The amount of algebra required can be reduced by noting that the gradients are all of the form

$$\frac{2}{\sigma^2} \sum_{m=1}^M \text{Re}\{\mathbf{F}^*(m)\mathbf{n}(m)\} \quad (\text{A.41})$$

where $\mathbf{F}(m)$ is some $QN \times K$ complex matrix. The general expression for a block in the information matrix is then

$$\begin{aligned}
& \mathbb{E} \left[\frac{2}{\sigma^2} \sum_{m=1}^M \operatorname{Re} \{ \mathbf{F}_1^*(m) \mathbf{n}(m) \} \left(\frac{2}{\sigma^2} \sum_{k=1}^M \operatorname{Re} \{ \mathbf{F}_2^*(k) \mathbf{n}(k) \} \right)^T \right] \\
&= \frac{4}{\sigma^4} \sum_{m=1}^M \sum_{k=1}^M \frac{1}{2} \mathbb{E} \left[\operatorname{Re} \{ \mathbf{F}_1^*(m) \mathbf{n}(m) (\mathbf{F}_2^*(k) \mathbf{n}(k))^T \} \right. \\
&\quad \left. + \operatorname{Re} \{ \mathbf{F}_1^*(m) \mathbf{n}(m) (\mathbf{F}_2^*(k) \mathbf{n}(k))^* \} \right] \\
&= \frac{2}{\sigma^4} \sum_{m=1}^M \sum_{k=1}^M \operatorname{Re} \{ \mathbf{F}_1^*(m) (\sigma^2 \mathbf{I}_{QN} \delta_K(k-m)) \mathbf{F}_2(k) \} \\
&= \frac{2}{\sigma^2} \sum_{m=1}^M \operatorname{Re} \{ \mathbf{F}_1^*(m) \mathbf{F}_2(m) \} \tag{A.42}
\end{aligned}$$

The blocks of $\mathbf{J}$, as found in (4.5)–(4.10), now follows by identifying the corresponding $\mathbf{F}(m)$ in (A.21), (A.24), and (A.33), and substituting the result into (A.42).

It appears to be difficult to simplify the expression for the information matrix any further. However, as the observation period becomes large, i.e., as $M \rightarrow \infty$, some simplifications can be made. The blocks of the information matrix in (4.4) are of the form

$$\operatorname{Re} \left\{ \sum_{m=1}^M \mathbf{Z}^*(m) \mathbf{H} \mathbf{Z}^*(m) \right\} \tag{A.43}$$

where $\mathbf{H}$ is some $2K \times 2K$ complex matrix. This can be exploited as follows. The (i, j) th element of $\mathbf{Z}^*(m) \mathbf{H} \mathbf{Z}(m)$ is $\mathbf{z}_i^* \mathbf{H} \mathbf{z}_j$ where $\mathbf{z}_k$ is the k th column of $\mathbf{Z}(m)$,

$$\mathbf{z}_k = [0 \quad \cdots \quad 0 \quad z_{2k-1}(m) \quad z_{2k}(m) \quad 0 \quad \cdots \quad 0]^T \tag{A.44}$$

Thus,

$$\mathbf{z}_i^* \mathbf{H} \mathbf{z}_j = [z_{2i-1}(m) \quad z_{2i}(m)] \mathbf{H}_{ij} \begin{bmatrix} z_{2j-1}(m) \\ z_{2j}(m) \end{bmatrix} \tag{A.45}$$

where $\mathbf{H}_{ij} \in \mathbb{C}^{2 \times 2}$ is defined as

$$\mathbf{H}_{ij} = \begin{bmatrix} \mathbf{H}(2i-1, 2j-1) & \mathbf{H}(2i-1, 2j) \\ \mathbf{H}(2i, 2j-1) & \mathbf{H}(2i, 2j) \end{bmatrix} \tag{A.46}$$

Expanding (A.45) yields

$$\begin{aligned} \mathbf{z}_i^* \mathbf{H} \mathbf{z}_j &= z_{2i-1}(m) z_{2j-1}(m) \mathbf{H}_{ij}(1, 1) + z_{2i}(m) z_{2j-1}(m) \mathbf{H}_{ij}(2, 1) \\ &\quad + z_{2i-1}(m) z_{2j}(m) \mathbf{H}_{ij}(1, 2) + z_{2i}(m) z_{2j}(m) \mathbf{H}_{ij}(2, 2) \end{aligned} \quad (\text{A.47})$$

As $M \rightarrow \infty$,

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M z_p(m) z_q(m) = \mathbb{E}[z_p(m) z_q(m)] \quad (\text{A.48})$$

and, since $d_i(m)$ and $d_j(m)$ are zero-mean and independent if $i \neq j$ and since $d_i(m)$ and $d_i(m-1)$ are independent,

$$\mathbb{E}[z_{2i-1}(m) z_{2j-1}(m)] = \frac{1}{2} \delta_K(i - j) \quad (\text{A.49})$$

$$\mathbb{E}[z_{2i}(m) z_{2j}(m)] = \frac{1}{2} \delta_K(i - j) \quad (\text{A.50})$$

$$\mathbb{E}[z_{2i-1}(m) z_{2j}(m)] = 0 \quad (\text{A.51})$$

Now,

$$\lim_{M \rightarrow \infty} \sum_{m=1}^M \mathbf{z}_i^* \mathbf{H} \mathbf{z}_j = \frac{M}{2} (\mathbf{H}_{ij}(1, 1) + \mathbf{H}_{ij}(2, 2)) \delta_K(i - j) \quad (\text{A.52})$$

It is not hard to see that

$$\lim_{M \rightarrow \infty} \sum_{m=1}^M \mathbf{Z}^*(m) \mathbf{H} \mathbf{Z}(m) = \frac{M}{2} \mathbf{G}^* \text{diag}(\mathbf{H}) \mathbf{G} \quad (\text{A.53})$$

where $\mathbf{G}$ is defined as in (4.20). In other words, (A.53) is a diagonal $K \times K$ matrix whose k th diagonal element is $(\mathbf{H}(2k-1, 2k-1) + \mathbf{H}(2k, 2k))M/2$. Note that if $\mathbf{X}, \mathbf{B} \in \mathbb{C}^{2K \times 2K}$ and if $\mathbf{B}$ is diagonal such that $\mathbf{B}(2k-1, 2k-1) = \mathbf{B}(2k, 2k)$, then

$$\mathbf{G}^* \text{diag}(\mathbf{B}^* \mathbf{X} \mathbf{B}) \mathbf{G} = \mathbf{G}^* \text{diag}(\mathbf{X}) \mathbf{G} \frac{1}{2} \mathbf{G}^* \mathbf{B}^* \mathbf{B} \mathbf{G} \quad (\text{A.54})$$

This is exactly the form of the blocks of $\mathbf{J}$. For instance, consider

$$\mathbf{J}_{\gamma\gamma, \infty} = \lim_{M \rightarrow \infty} \mathbf{J}_{\gamma\gamma} = \lim_{M \rightarrow \infty} \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \left\{ \mathbf{Z}^*(m) \mathbf{B}_\gamma^* \mathbf{A}^* \mathbf{A} \mathbf{B}_\gamma \mathbf{Z}(m) \right\}$$

$$\begin{aligned}
&= \frac{M}{\sigma^2} \operatorname{Re} \left\{ \mathbf{G}^* \operatorname{diag}(\mathbf{B}_\gamma^* \mathbf{A}^* \mathbf{A} \mathbf{B}_\gamma) \mathbf{G} \right\} \\
&= \frac{M}{\sigma^2} \operatorname{Re} \left\{ \mathbf{G}^* \operatorname{diag}(\mathbf{A}^* \mathbf{A}) \mathbf{G} \frac{1}{2} \mathbf{G}^* \mathbf{B}_\gamma^* \mathbf{B}_\gamma \mathbf{G} \right\} \\
&= \frac{M}{\sigma^2} \operatorname{Re} \left\{ \mathbf{G}^* \operatorname{diag}(\mathbf{A}^* \mathbf{A}) \mathbf{G} \right\}
\end{aligned} \tag{A.55}$$

since $\mathbf{B}_\gamma^* \mathbf{B}_\gamma = \mathbf{I}_{2K}$ and $\mathbf{G}^* \mathbf{I}_{2K} \mathbf{G} = 2\mathbf{I}_K$. Furthermore, $\mathbf{B}_\gamma^* \mathbf{B}_\theta = j|\mathbf{B}|$, $\mathbf{B}_\gamma^* \mathbf{B} = |\mathbf{B}|$, $\mathbf{B}_\theta^* \mathbf{B}_\theta = |\mathbf{B}|^2$, $\mathbf{B}_\theta^* \mathbf{B} = -j|\mathbf{B}|^2$, and $\mathbf{B}^* \mathbf{B} = |\mathbf{B}|^2$, where we by $|\mathbf{B}|$ denote the matrix whose elements are $|\mathbf{B}|(i, j) = |\mathbf{B}(i, j)|$. The blocks of $\mathbf{J}$ as $M \rightarrow \infty$ can now be written as

$$\mathbf{J}_{\gamma\theta, \infty} = \lim_{M \rightarrow \infty} \mathbf{J}_{\gamma\theta} = \mathbf{0} \tag{A.56}$$

$$\begin{aligned}
\mathbf{J}_{\gamma\tau, \infty} &= \lim_{M \rightarrow \infty} \mathbf{J}_{\gamma\tau} = \frac{M}{\sigma^2} \mathbf{G}^* \operatorname{diag}(\mathbf{A}^* \Psi) \mathbf{G} \frac{1}{2} \mathbf{G}^* |\mathbf{B}| \mathbf{G} \\
&= \frac{M}{\sigma^2} \mathbf{G}^* \operatorname{diag}(\mathbf{A}^* \Psi) \mathbf{G} \sqrt{\mathbf{P}}
\end{aligned} \tag{A.57}$$

$$\begin{aligned}
\mathbf{J}_{\theta\theta, \infty} &= \lim_{M \rightarrow \infty} \mathbf{J}_{\theta\theta} = \frac{M}{\sigma^2} \mathbf{G}^* \operatorname{diag}(\mathbf{A}^* \mathbf{A}) \mathbf{G} \frac{1}{2} \mathbf{G}^* |\mathbf{B}|^2 \mathbf{G} \\
&= \frac{M}{\sigma^2} \mathbf{G}^* \operatorname{diag}(\mathbf{A}^* \mathbf{A}) \mathbf{G} \mathbf{P}
\end{aligned} \tag{A.58}$$

$$\mathbf{J}_{\theta\tau, \infty} = \lim_{M \rightarrow \infty} \mathbf{J}_{\theta\tau} = \mathbf{0} \tag{A.59}$$

$$\begin{aligned}
\mathbf{J}_{\tau\tau, \infty} &= \lim_{M \rightarrow \infty} \mathbf{J}_{\tau\tau} = \frac{M}{\sigma^2} \mathbf{G}^* \operatorname{diag}(\Psi^* \Psi) \mathbf{G} \frac{1}{2} \mathbf{G}^* |\mathbf{B}|^2 \mathbf{G} \\
&= \frac{M}{\sigma^2} \mathbf{G}^* \operatorname{diag}(\Psi^* \Psi) \mathbf{G} \mathbf{P}
\end{aligned} \tag{A.60}$$

where the power matrix $\mathbf{P} \in \mathbb{R}^{K \times K}$ is defined by (4.21). The expression (4.15) for asymptotic information matrix now follows.

A.2 CRB for Fading Channels

We will now repeat the CRB derivation for the fading channel model as described in Chapter 6.

A.2.1 Differentiating the Log-Likelihood Function

It is easy to show that the log-likelihood function (conditioned on $\mathbf{z}$) of $\mathbf{r}$ with respect to $\boldsymbol{\phi}$ is

$$\ln L(\mathbf{r}) = -MQN \ln \sigma^2 - \frac{1}{\sigma^2} \sum_{m=1}^M \|\mathbf{r}(m) - \mathbf{A}\mathbf{s}(m)\|^2 \quad (\text{A.61})$$

Computing $\partial \ln L(\mathbf{r}) / \partial \sigma^2$

We start by differentiating the log-likelihood function with respect to σ^2 ,

$$\frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} = -MQN \frac{1}{\sigma^2} + \frac{1}{\sigma^4} \sum_{m=1}^M \|\mathbf{n}(m)\|^2 \quad (\text{A.62})$$

Computing $\partial \ln L(\mathbf{r}) / \partial \tilde{\boldsymbol{\beta}}(m)$

Consider

$$\frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} = -\frac{1}{\sigma^2} \frac{\partial}{\partial \tilde{\boldsymbol{\beta}}(m)} \|\mathbf{r}(m) - \mathbf{A}\mathbf{s}(m)\|^2 \quad (\text{A.63})$$

With the above Results 1, 2, and 3 in mind, we can easily show that

$$\begin{aligned} \frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} &= \frac{1}{\sigma^2} \left[\left(\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}(m)} \right)^T \mathbf{A}^T \bar{\mathbf{n}}(m) + \left(\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}(m)} \right)^* \mathbf{A}^* \mathbf{n}(m) \right] \\ &= \frac{1}{\sigma^2} 2 \operatorname{Re} \left[\left(\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}(m)} \right)^* \mathbf{A}^* \mathbf{n}(m) \right] \end{aligned} \quad (\text{A.64})$$

where

$$\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}(m)} = \begin{bmatrix} \frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}_1(m)} & \frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}_2(m)} & \cdots & \frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}_K(m)} \end{bmatrix} \in \mathbb{C}^{2R \times R} \quad (\text{A.65})$$

$$\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}_k(m)} = \begin{bmatrix} \frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}_{k,1}(m)} & \frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}_{k,2}(m)} & \cdots & \frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}_{k,R_k}(m)} \end{bmatrix} \in \mathbb{C}^{2R \times R_k} \quad (\text{A.66})$$

$$\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}_{k,r}(m)} = \begin{bmatrix} 0 & \cdots & 0 & \frac{\partial s_{k,2r-1}(m)}{\partial \tilde{\boldsymbol{\beta}}_{k,r}(m)} & \frac{\partial s_{k,2r}(m)}{\partial \tilde{\boldsymbol{\beta}}_{k,r}(m)} & 0 & \cdots & 0 \end{bmatrix}^T \in \mathbb{C}^{2R_k} \quad (\text{A.67})$$

$\underbrace{\hspace{1.5cm}}_{2(r-1) \text{ zeros}}$

Recall that

$$s_{k,2r-1}(m) = \beta_{k,r}(m) z_{2k-1}(m) \quad (\text{A.68})$$

$$s_{k,2r}(m) = \beta_{k,r}(m) z_{2k}(m) \quad (\text{A.69})$$

and therefore

$$\frac{\partial s_{k,2r-1}(m)}{\partial \tilde{\beta}_{k,r}(m)} = \frac{\partial}{\partial \tilde{\beta}_{k,r}(m)} [\tilde{\beta}_{k,r}(m) + j \tilde{\beta}_{k,r}(m)] z_{2k-1}(m) = z_{2k-1}(m) \quad (\text{A.70})$$

and

$$\frac{\partial s_{k,2r}(m)}{\partial \tilde{\beta}_{k,r}(m)} = z_{2k}(m) \quad (\text{A.71})$$

If we define $\mathbf{Z}_k(m) \in \{-1, 0, 1\}^{2R_k \times R_k}$ as

$$(\mathbf{Z}_k(m))(i, j) = \begin{cases} z_{2k-1}(m) & \text{if } i = 2j - 1 \\ z_{2k}(m) & \text{if } i = 2j \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.72})$$

then

$$\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}_k(m)} = \begin{bmatrix} \mathbf{0} \\ \mathbf{Z}_k(m) \\ \mathbf{0} \end{bmatrix}. \quad (\text{A.73})$$

where there are $\sum_{l=1}^{k-1} 2R_l$ leading zero rows in $\partial \mathbf{s}(m)/\partial \tilde{\boldsymbol{\beta}}_k(m)$.

We can now easily form $\partial \mathbf{s}(m)/\partial \tilde{\boldsymbol{\beta}}(m)$ as

$$\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}(m)} = \text{diag}(\mathbf{Z}_1(m), \mathbf{Z}_2(m), \dots, \mathbf{Z}_K(m)) = \mathbf{Z}(m) \quad (\text{A.74})$$

Computing $\partial \ln L(\mathbf{r})/\partial \tilde{\boldsymbol{\beta}}(m)$

Similarly as for $\tilde{\boldsymbol{\beta}}(m)$, it can shown that

$$\frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} = \frac{1}{\sigma^2} 2 \text{Re} \left[\left(\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}(m)} \right)^* \mathbf{A}^* \mathbf{n}(m) \right] \quad (\text{A.75})$$

and

$$\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}(m)} = j \frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}(m)} \quad (\text{A.76})$$

Thus,

$$\begin{aligned} \frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} &= \frac{1}{\sigma^2} 2 \text{Re} \left[-j \left(\frac{\partial \mathbf{s}(m)}{\partial \tilde{\boldsymbol{\beta}}(m)} \right)^* \mathbf{A}^* \mathbf{n}(m) \right] \\ &= \frac{2}{\sigma^2} \text{Im} [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{n}(m)] \end{aligned} \quad (\text{A.77})$$

Computing $\partial \ln L(\mathbf{r})/\partial \boldsymbol{\tau}$

We will now turn our attention to the derivative of $\ln L(\mathbf{r})$ with respect to $\boldsymbol{\tau}$.

Using Results 2 and 3, we can show that

$$\begin{aligned} \frac{\partial \ln L(\mathbf{r})}{\partial \tau_{k,r}} &= \frac{1}{\sigma^2} \sum_{m=1}^M \mathbf{s}^T(m) \left(\frac{\partial \mathbf{A}}{\partial \tau_{k,r}} \right)^T \bar{\mathbf{n}}(m) + \mathbf{s}^*(m) \left(\frac{\partial \mathbf{A}}{\partial \tau_{k,r}} \right)^* \mathbf{n}(m) \\ &= \frac{2}{\sigma^2} \sum_{m=1}^M \operatorname{Re} \left[\mathbf{s}^*(m) \left(\frac{\partial \mathbf{A}}{\partial \tau_{k,r}} \right)^* \mathbf{n}(m) \right] \end{aligned} \quad (\text{A.78})$$

Consider

$$\frac{\partial \mathbf{A}}{\partial \tau_{k,r}} = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ \underbrace{\hspace{1.5cm}}_{k-1 \text{ matrices}} & \frac{\partial \mathbf{A}_k}{\partial \tau_{k,r}} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{QN \times 2R} \quad (\text{A.79})$$

$$\frac{\partial \mathbf{A}_k}{\partial \tau_{k,r}} = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ \underbrace{\hspace{1.5cm}}_{2(r-1) \text{ columns}} & \frac{\partial \mathbf{a}_{k,2r-1}}{\partial \tau_{k,r}} & \frac{\partial \mathbf{a}_{k,2r}}{\partial \tau_{k,r}} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{QN \times 2R_k} \quad (\text{A.80})$$

We define $\boldsymbol{\psi}_{k,2r-1}, \boldsymbol{\psi}_{k,2r} \in \mathbb{R}^{QN}$ as

$$\boldsymbol{\psi}_{k,2r-1} = \frac{\partial \mathbf{a}_{k,2r-1}}{\partial \tau_{k,r}}, \quad \boldsymbol{\psi}_{k,2r} = \frac{\partial \mathbf{a}_{k,2r}}{\partial \tau_{k,r}} \quad (\text{A.81})$$

It is easy to see that

$$\mathbf{s}^*(m) \left(\frac{\partial \mathbf{A}}{\partial \tau_{k,r}} \right)^* = \begin{bmatrix} \bar{s}_{k,2r-1} & \bar{s}_{k,2r} \end{bmatrix} \begin{bmatrix} \boldsymbol{\psi}_{k,2r-1}^* \\ \boldsymbol{\psi}_{k,2r}^* \end{bmatrix} \quad (\text{A.82})$$

and thus

$$\mathbf{s}^*(m) \left(\frac{\partial \mathbf{A}}{\partial \tau_{k,r}} \right)^* = \boldsymbol{\Gamma}_k^*(m) \boldsymbol{\Psi}_k^* \quad (\text{A.83})$$

where we have defined $\boldsymbol{\Gamma}_k(m) \in \mathbb{C}^{2R_k \times R_k}$ and $\boldsymbol{\Psi}_k \in \mathbb{R}^{QN \times 2R_k}$ as

$$(\boldsymbol{\Gamma}_k(m))(i, j) = \begin{cases} s_{k,i}(m) & \text{if } i \in \{2j-1, 2j\} \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.84})$$

$$\boldsymbol{\Psi}_k = [\boldsymbol{\psi}_{k,1} \quad \boldsymbol{\psi}_{k,2} \quad \cdots \quad \boldsymbol{\psi}_{k,2R_k}] \quad (\text{A.85})$$

Finally, we can form

$$\frac{\partial \ln L(\mathbf{r})}{\partial \boldsymbol{\tau}} = \frac{2}{\sigma^2} \sum_{m=1}^M \operatorname{Re} [\boldsymbol{\Gamma}^*(m) \boldsymbol{\Psi}^* \mathbf{n}(m)] \quad (\text{A.86})$$

where

$$\mathbf{\Gamma}(m) = \text{diag}(\mathbf{\Gamma}_1(m), \mathbf{\Gamma}_2(m), \dots, \mathbf{\Gamma}_K(m)) \in \mathbb{C}^{2R \times R} \quad (\text{A.87})$$

$$\mathbf{\Psi} = [\mathbf{\Psi}_1 \quad \mathbf{\Psi}_2 \quad \dots \quad \mathbf{\Psi}_K] \in \mathbb{R}^{QN \times 2R} \quad (\text{A.88})$$

Summary

As shown above, the derivatives of the log-likelihood function can be written as

$$\frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} = -\frac{MQN}{\sigma^2} + \frac{1}{\sigma^4} \sum_{m=1}^M \|\mathbf{n}(m)\|^2 \quad (\text{A.89})$$

$$\frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} = \frac{2}{\sigma^2} \text{Re} [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{n}(m)] \quad (\text{A.90})$$

$$\frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} = \frac{2}{\sigma^2} \text{Im} [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{n}(m)] \quad (\text{A.91})$$

$$\frac{\partial \ln L(\mathbf{r})}{\partial \boldsymbol{\tau}} = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} [\mathbf{\Gamma}^*(m) \mathbf{\Psi}^* \mathbf{n}(m)] \quad (\text{A.92})$$

A.2.2 Forming the Fisher Information Matrix

We form the blocks of $\mathbf{J}$ by taking the expected value of all possible outer products of the matrices (A.89), (A.90), (A.91), and (A.92).

Computing $\mathbb{E}[(\partial \ln L(\mathbf{r}) / \partial \sigma^2)(\cdot)^T]$

The (1, 1) block in $\mathbf{J}$ is

$$\begin{aligned} \mathbb{E} \left[\frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} \left(\frac{\partial \ln L}{\partial \sigma^2} \right)^T \right] &= \frac{(MQN)^2}{\sigma^4} - 2 \frac{MQN}{\sigma^2} \frac{1}{\sigma^4} \sum_{m=1}^M \mathbb{E}[\mathbf{n}^*(m) \mathbf{n}(m)] \\ &\quad + \frac{1}{\sigma^8} \sum_{m=1}^M \sum_{p=1}^M \mathbb{E}[\mathbf{n}^*(m) \mathbf{n}(m) \mathbf{n}^*(p) \mathbf{n}(p)] \\ &= \frac{(MQN)^2}{\sigma^4} - 2 \frac{(MQN)^2}{\sigma^4} \\ &\quad + \frac{1}{\sigma^8} [(M^2 - M)(QN)^2 \sigma^4 + MQN(QN + 1) \sigma^4] \\ &= \frac{MQN}{\sigma^4} \end{aligned} \quad (\text{A.93})$$

By using Result 5, it can be shown $\partial \ln L / \partial \sigma^2$ is uncorrelated with all other derivatives. Thus the $(1, n)$ blocks (and, therefore, the $(n, 1)$ blocks) of $\mathbf{J}$ will be zero.

Computing $E[(\partial \ln L(\mathbf{r}) / \partial \tilde{\boldsymbol{\beta}}(m))(\partial \ln L(\mathbf{r}) / \partial \tilde{\boldsymbol{\beta}}(p))^T]$

Some straightforward calculations using Result 6 and the fact that $E[\mathbf{n}(m)\mathbf{n}^T(p)] = \mathbf{0}$ yield

$$\begin{aligned} E \left[\frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} \left(\frac{\partial \ln L}{\tilde{\boldsymbol{\beta}}(p)} \right)^T \right] &= \frac{4}{\sigma^4} E \left\{ \text{Re} [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{n}(m)] \text{Re}^T [\mathbf{Z}^*(p) \mathbf{A}^* \mathbf{n}(p)] \right\} \\ &= \frac{4}{\sigma^4} \frac{1}{2} \text{Re} [\mathbf{Z}^*(m) \mathbf{A}^* E\{\mathbf{n}(m)\mathbf{n}^*(p)\} \mathbf{A} \mathbf{Z}(p)] \\ &= \frac{2}{\sigma^2} \text{Re} [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(p)] \delta_K(m - p) \end{aligned} \quad (\text{A.94})$$

Computing $E[(\partial \ln L(\mathbf{r}) / \partial \tilde{\boldsymbol{\beta}}(m))(\partial \ln L(\mathbf{r}) / \partial \tilde{\boldsymbol{\beta}}(p))^T]$

Similar to above,

$$\begin{aligned} E \left[\frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} \left(\frac{\partial \ln L}{\tilde{\boldsymbol{\beta}}(p)} \right)^T \right] &= \frac{4}{\sigma^4} E \left\{ \text{Re} [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{n}(m)] \text{Im}^T [\mathbf{Z}^*(p) \mathbf{A}^* \mathbf{n}(p)] \right\} \\ &= -\frac{4}{\sigma^4} \frac{1}{2} \text{Im} [\mathbf{Z}^*(m) \mathbf{A}^* E\{\mathbf{n}(m)\mathbf{n}^*(p)\} \mathbf{A} \mathbf{Z}(p)] \\ &= -\frac{2}{\sigma^2} \text{Im} [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(p)] \delta_K(m - p) \end{aligned} \quad (\text{A.95})$$

Computing $E[(\partial \ln L(\mathbf{r}) / \partial \tilde{\boldsymbol{\beta}}(m))(\partial \ln L(\mathbf{r}) / \partial \tilde{\boldsymbol{\beta}}(p))^T]$

Again, as above,

$$\begin{aligned} E \left[\frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} \left(\frac{\partial \ln L}{\tilde{\boldsymbol{\beta}}(p)} \right)^T \right] &= \frac{4}{\sigma^4} E \left\{ \text{Im} [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{n}(m)] \text{Im}^T [\mathbf{Z}^*(p) \mathbf{A}^* \mathbf{n}(p)] \right\} \\ &= \frac{4}{\sigma^4} \frac{1}{2} \text{Re} [\mathbf{Z}^*(m) \mathbf{A}^* E\{\mathbf{n}(m)\mathbf{n}^*(p)\} \mathbf{A} \mathbf{Z}(p)] \\ &= \frac{2}{\sigma^2} \text{Re} [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(p)] \delta_K(m - p) \\ &= E \left[\frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} \left(\frac{\partial \ln L(\mathbf{r})}{\tilde{\boldsymbol{\beta}}(p)} \right)^T \right] \end{aligned} \quad (\text{A.96})$$

Computing $E[(\partial \ln L(\mathbf{r})/\partial \tilde{\boldsymbol{\beta}}(m))(\partial \ln L(\mathbf{r})/\partial \boldsymbol{\tau})^T]$

Straightforward computations show

$$\begin{aligned}
 E \left[\frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} \left(\frac{\partial \ln L}{\partial \boldsymbol{\tau}} \right)^T \right] &= \frac{4}{\sigma^4} E \left\{ \text{Re} [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{n}(m)] \text{Re}^T \left[\sum_{p=1}^M \boldsymbol{\Gamma}^*(p) \boldsymbol{\Psi}^* \mathbf{n}(p) \right] \right\} \\
 &= \frac{4}{\sigma^4} \frac{1}{2} \sum_{p=1}^M \text{Re} [\mathbf{Z}^*(m) \mathbf{A}^* E\{\mathbf{n}(m) \mathbf{n}^*(p)\} \boldsymbol{\Psi} \boldsymbol{\Gamma}(p)] \\
 &= \frac{2}{\sigma^2} \text{Re}[\mathbf{Z}^*(m) \mathbf{A}^* \boldsymbol{\Psi} \boldsymbol{\Gamma}(m)] \quad (\text{A.97})
 \end{aligned}$$

Computing $E[(\partial \ln L(\mathbf{r})/\partial \tilde{\boldsymbol{\beta}}(m))(\partial \ln L(\mathbf{r})/\partial \boldsymbol{\tau})^T]$

Straightforward computations show

$$\begin{aligned}
 E \left[\frac{\partial \ln L(\mathbf{r})}{\partial \tilde{\boldsymbol{\beta}}(m)} \left(\frac{\partial \ln L}{\partial \boldsymbol{\tau}} \right)^T \right] &= E \left[\frac{\partial \ln L(\mathbf{r})}{\partial \boldsymbol{\tau}} \left(\frac{\partial \ln L}{\partial \tilde{\boldsymbol{\beta}}(m)} \right)^T \right]^T \\
 &= \frac{4}{\sigma^4} E \left\{ \text{Re} \left[\sum_{p=1}^M \boldsymbol{\Gamma}^*(p) \boldsymbol{\Psi}^* \mathbf{n}(p) \right] \text{Im}^T [\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{n}(m)] \right\}^T \\
 &= -\frac{4}{\sigma^4} \frac{1}{2} \sum_{p=1}^M \text{Im}^T [\boldsymbol{\Gamma}^*(p) \boldsymbol{\Psi}^* E\{\mathbf{n}(m) \mathbf{n}^*(p)\} \mathbf{A} \mathbf{Z}(m)] \\
 &= \frac{2}{\sigma^2} \text{Im}[\mathbf{Z}^*(m) \mathbf{A}^* \boldsymbol{\Psi} \boldsymbol{\Gamma}(m)] \quad (\text{A.98})
 \end{aligned}$$

Computing $E[(\partial \ln L(\mathbf{r})/\partial \boldsymbol{\tau})(\partial \ln L(\mathbf{r})/\partial \boldsymbol{\tau})^T]$

The last block of the information matrix is

$$\begin{aligned}
 E \left[\frac{\partial \ln L(\mathbf{r})}{\partial \boldsymbol{\tau}} \left(\frac{\partial \ln L}{\partial \boldsymbol{\tau}} \right)^T \right] &= \frac{4}{\sigma^4} E \left\{ \text{Re} \left[\sum_{m=1}^M \boldsymbol{\Gamma}^*(m) \boldsymbol{\Psi}^* \mathbf{n}(m) \right] \text{Re}^T \left[\sum_{p=1}^M \boldsymbol{\Gamma}^*(p) \boldsymbol{\Psi}^* \mathbf{n}(p) \right] \right\} \\
 &= \frac{4}{\sigma^4} \frac{1}{2} \sum_{m=1}^M \sum_{p=1}^M \text{Re} [\boldsymbol{\Gamma}^*(m) \boldsymbol{\Psi}^* E\{\mathbf{n}(m) \mathbf{n}^*(p)\} \boldsymbol{\Psi} \boldsymbol{\Gamma}(p)] \\
 &= \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re}[\boldsymbol{\Gamma}^*(m) \boldsymbol{\Psi}^* \boldsymbol{\Psi} \boldsymbol{\Gamma}(m)] \quad (\text{A.99})
 \end{aligned}$$

Assembling the Fisher information matrix

We are now ready to form $\mathbf{J}$. The information matrix is partitioned as

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\sigma\sigma} & \mathbf{J}_{\sigma s} & \mathbf{J}_{\sigma\tau} \\ \mathbf{J}_{\sigma s}^T & \mathbf{J}_{ss} & \mathbf{J}_{s\tau} \\ \mathbf{J}_{\sigma\tau}^T & \mathbf{J}_{s\tau}^T & \mathbf{J}_{\tau\tau} \end{bmatrix} \quad (\text{A.100})$$

where

$$\mathbf{J}_{\sigma\sigma} = \mathbb{E} \left[\frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} \left(\frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} \right)^T \right] = \frac{MQN}{\sigma^4} \quad (\text{A.101})$$

$$\mathbf{J}_{\sigma s} = \mathbb{E} \left[\frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} \left(\frac{\partial \ln L(\mathbf{r})}{\partial \beta} \right)^T \right] = \mathbf{0} \quad (\text{A.102})$$

$$\mathbf{J}_{\sigma\tau} = \mathbb{E} \left[\frac{\partial \ln L(\mathbf{r})}{\partial \sigma^2} \left(\frac{\partial \ln L(\mathbf{r})}{\partial \tau} \right)^T \right] = \mathbf{0} \quad (\text{A.103})$$

$$\mathbf{J}_{\tau\tau} = \mathbb{E} \left[\frac{\partial \ln L}{\partial \tau} \left(\frac{\partial \ln L}{\partial \tau} \right)^T \right] = \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re}[\Gamma^*(m) \Psi^* \Psi \Gamma(m)] \quad (\text{A.104})$$

and where $\mathbf{J}_{ss}$ and $\mathbf{J}_{s\tau}$ are defined as

$$\mathbf{J}_{ss} = \mathbb{E} \left[\frac{\partial \ln L(\mathbf{r})}{\partial \beta} \left(\frac{\partial \ln L(\mathbf{r})}{\partial \beta} \right)^T \right] \quad (\text{A.105})$$

$$\mathbf{J}_{s\tau} = \mathbb{E} \left[\frac{\partial \ln L(\mathbf{r})}{\partial \beta} \left(\frac{\partial \ln L(\mathbf{r})}{\partial \tau} \right)^T \right] \quad (\text{A.106})$$

From (A.94), (A.95), and (A.96), we note that the derivative of $\ln L(\mathbf{r})$ with respect to $\beta(m)$ and $\beta(p)$ are uncorrelated when $m \neq p$. $\mathbf{J}_{ss}$ will therefore be block diagonal

$$\mathbf{J}_{ss} = \text{diag}(\mathbf{J}_{ss}(1), \mathbf{J}_{ss}(2), \dots, \mathbf{J}_{ss}(M)) \quad (\text{A.107})$$

where the m th block is

$$\mathbf{J}_{ss}(m) = \frac{2}{\sigma^2} \begin{bmatrix} \text{Re}[\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(m)] & -\text{Im}[\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(m)] \\ -\text{Im}^T[\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(m)] & \text{Re}[\mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(m)] \end{bmatrix} \quad (\text{A.108})$$

Furthermore, from (A.97) and (A.98) we see that

$$\mathbf{J}_{s\tau} = [\mathbf{J}_{s\tau}^T(1) \quad \mathbf{J}_{s\tau}^T(2) \quad \dots \quad \mathbf{J}_{s\tau}^T(M)]^T \quad (\text{A.109})$$

where

$$\mathbf{J}_{s\tau}(m) = \frac{2}{\sigma^2} \begin{bmatrix} \text{Re}[\mathbf{Z}^*(m)\mathbf{A}^*\Psi\Gamma(m)] \\ \text{Im}[\mathbf{Z}^*(m)\mathbf{A}^*\Psi\Gamma(m)] \end{bmatrix} \quad (\text{A.110})$$

In summary, we can write the information matrix as

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\sigma\sigma} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{ss} & \mathbf{J}_{s\tau} \\ \mathbf{0} & \mathbf{J}_{s\tau}^T & \mathbf{J}_{\tau\tau} \end{bmatrix} \quad (\text{A.111})$$

A.2.3 Inverting the Information Matrix

The CRB matrix is found as the inverse of the Fisher information matrix. We are mostly interested in bounding the performance of a propagation delay estimator. The bound on the covariance matrix of such an estimator is found as the lower right hand side $R \times R$ block of the CRB matrix. In other words, we can write the CRB matrix as

$$\mathbf{J}^{-1} = \begin{bmatrix} \text{CRB}(\sigma^2) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \text{CRB}(s) & \times \\ \mathbf{0} & \times & \text{CRB}(\tau) \end{bmatrix} \quad (\text{A.112})$$

where $\times$ denotes matrices that are of no particular interest to us and where

$$\text{CRB}(\tau) \leq \mathbb{E}[(\hat{\tau} - \tau)(\hat{\tau} - \tau)^T] \quad (\text{A.113})$$

for any unbiased estimator $\hat{\tau}$ of the propagation delay vector τ .

Invoking a standard result on the inverse of a block matrix (see [40], Lemma 1, page 53), we see from (A.111) that

$$\text{CRB}^{-1}(\tau) = \mathbf{J}_{\tau\tau} - \mathbf{J}_{s\tau}^T \mathbf{J}_{ss}^{-1} \mathbf{J}_{s\tau} \quad (\text{A.114})$$

To compute the last term of this equation, we first make a few definitions for notational convenience. Let

$$\mathbf{H}(m) = \frac{2}{\sigma^2} \mathbf{Z}^*(m) \mathbf{A}^* \mathbf{A} \mathbf{Z}(m) \quad (\text{A.115})$$

$$\mathbf{G}(m) = \mathbf{H}^{-1}(m) \quad (\text{A.116})$$

$$\Delta(m) = \frac{2}{\sigma^2} \mathbf{Z}^*(m) \mathbf{A}^* \Psi \Gamma(m) \quad (\text{A.117})$$

Note, since $\mathbf{H}(m)$ is hermitian, that $\text{Im}[\mathbf{H}(m)]$ must be skew symmetric,

$$\text{Im}^T[\mathbf{H}(m)] = -\text{Im}[\mathbf{H}(m)] \quad (\text{A.118})$$

With this in mind, we can write $\mathbf{J}_{ss}(m)$ as

$$\mathbf{J}_{ss}(m) = \begin{bmatrix} \tilde{\mathbf{H}}(m) & -\tilde{\mathbf{H}}(m) \\ \tilde{\mathbf{H}}(m) & \tilde{\mathbf{H}}(m) \end{bmatrix} \quad (\text{A.119})$$

It is easy to show that

$$\mathbf{J}_{ss}^{-1}(m) = \begin{bmatrix} \tilde{\mathbf{G}}(m) & -\tilde{\mathbf{G}}(m) \\ \tilde{\mathbf{G}}(m) & \tilde{\mathbf{G}}(m) \end{bmatrix} \quad (\text{A.120})$$

Now,

$$\begin{aligned} \mathbf{J}_{s\tau}^T \mathbf{J}_{ss}^{-1} \mathbf{J}_{s\tau} &= \mathbf{J}_{s\tau}^T \begin{bmatrix} \mathbf{J}_{ss}^{-1}(1) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{ss}^{-1}(2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{J}_{ss}^{-1}(M) \end{bmatrix} \mathbf{J}_{s\tau} \\ &= \sum_{m=1}^M \mathbf{J}_{s\tau}^T(m) \mathbf{J}_{ss}^{-1}(m) \mathbf{J}_{s\tau}(m) \\ &= \sum_{m=1}^M \begin{bmatrix} \tilde{\Delta}^T(m) & \tilde{\Delta}^T(m) \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{G}}(m) & -\tilde{\mathbf{G}}(m) \\ \tilde{\mathbf{G}}(m) & \tilde{\mathbf{G}}(m) \end{bmatrix} \begin{bmatrix} \tilde{\Delta}(m) \\ \tilde{\Delta}(m) \end{bmatrix} \\ &= \sum_{m=1}^M \begin{bmatrix} \tilde{\Delta}^T(m) & \tilde{\Delta}^T(m) \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{G}}(m) \tilde{\Delta}(m) - \tilde{\mathbf{G}}(m) \tilde{\Delta}(m) \\ \tilde{\mathbf{G}}(m) \tilde{\Delta}(m) + \tilde{\mathbf{G}}(m) \tilde{\Delta}(m) \end{bmatrix} \\ &= \sum_{m=1}^M \begin{bmatrix} \tilde{\Delta}^T(m) & \tilde{\Delta}^T(m) \end{bmatrix} \begin{bmatrix} \text{Re}[\mathbf{G}(m) \Delta(m)] \\ \text{Im}[\mathbf{G}(m) \Delta(m)] \end{bmatrix} \\ &= \sum_{m=1}^M \tilde{\Delta}^T(m) \text{Re}[\mathbf{G}(m) \Delta(m)] + \tilde{\Delta}^T(m) \text{Im}[\mathbf{G}(m) \Delta(m)] \\ &= \sum_{m=1}^M \text{Re}[\Delta^*(m) \mathbf{G}(m) \Delta(m)] \quad (\text{A.121}) \end{aligned}$$

Finally, we can find the desired CRB as

$$\begin{aligned}
\text{CRB}^{-1}(\boldsymbol{\tau}) &= \mathbf{J}_{\boldsymbol{\tau}\boldsymbol{\tau}} - \sum_{m=1}^M \text{Re}[\boldsymbol{\Delta}^*(m)\mathbf{G}(m)\boldsymbol{\Delta}(m)] \\
&= \sum_{m=1}^M \text{Re} \left[\frac{2}{\sigma^2} \boldsymbol{\Gamma}^*(m)\boldsymbol{\Psi}^*\boldsymbol{\Psi}\boldsymbol{\Gamma}(m) - \boldsymbol{\Delta}^*(m)\mathbf{G}(m)\boldsymbol{\Delta}(m) \right] \\
&= \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \left[\boldsymbol{\Gamma}^*(m)\boldsymbol{\Psi}^*\boldsymbol{\Psi}\boldsymbol{\Gamma}(m) \right. \\
&\quad \left. - \boldsymbol{\Gamma}^*(m)\boldsymbol{\Psi}^*\mathbf{A}\mathbf{Z}(m)(\mathbf{Z}^*(m)\mathbf{A}^*\mathbf{A}\mathbf{Z}(m))^{-1}\mathbf{Z}^*(m)\mathbf{A}^*\boldsymbol{\Psi}\boldsymbol{\Gamma}(m) \right] \\
&= \frac{2}{\sigma^2} \sum_{m=1}^M \text{Re} \left[\boldsymbol{\Gamma}^*(m)\boldsymbol{\Psi}^*\mathbf{P}_{\mathbf{A}\mathbf{Z}(m)}^\perp\boldsymbol{\Psi}\boldsymbol{\Gamma}(m) \right] \tag{A.122}
\end{aligned}$$

where $\mathbf{P}_{\mathbf{A}\mathbf{Z}(m)}^\perp$ is the projection matrix onto $[\text{range}(\mathbf{A}\mathbf{Z}(m))]^\perp$

$$\mathbf{P}_{\mathbf{A}\mathbf{Z}(m)}^\perp = \mathbf{I}_{Q_N} - \mathbf{A}\mathbf{Z}(m)(\mathbf{Z}^*(m)\mathbf{A}^*\mathbf{A}\mathbf{Z}(m))^{-1}\mathbf{Z}^*(m)\mathbf{A}^* \tag{A.123}$$

APPENDIX B TYPOGRAPHICAL NOTES

This dissertation was typed and edited in Emacs version 19.25 using AUC TeX mode version 9.1i and was typeset by L^AT_EX2e version <1994/06/01> patch level 4. The key to symbols was compiled by Makeindex portable version 2.12 [26-May-1993]. The author is extremely grateful to the wonderful people that have developed these softwares and provided them free of charge to anyone on the planet.

The dissertation was produced in a single L^AT_EX2e run. An excerpt from the log-file is

Here is how much of TeX's memory you used:

2035 strings out of 3546

20909 string characters out of 51429

718225 words of memory out of 262141

4154 multiletter control sequences out of 9500

12878 words of font info for 45 fonts, out of 72000 for 255

14 hyphenation exceptions out of 607

30i,25n,24p,458b,513s stack positions out of 300i,40n,60p,3000b,4000s

Output written on thesis.dvi (168 pages, 545648 bytes).

The graduate school at the University of Florida has some very peculiar ideas on how a thesis should be typeset. As a result of the typesetting rules, the author was forced to spend several weeks on modifying the standard L^AT_EX2e book class into a terrible hack called `ufthesis.sty`. The source code for `ufthesis.sty` can be requested by mailing the author at `strom@elixir.e.kth.se`. However, the author will not accept any role as maintainer of this software. Moreover, he will not be able to answer any questions regarding its use.

REFERENCES

- [1] M. Abdulrahman, D. D. Falconer, and A. U. H. Sheikh. "Equalization for interference cancellation in spread spectrum multiple access systems." In *Proceedings IEEE Vehicular Technology Conference*, pp. 71–74, IEEE, Piscataway, New Jersey, 1992.
- [2] S. Andersson. *On Dimension Reduction in Sensor Array Signal Processing*. Ph.D. thesis, Linköping University, Dept. of Electrical Engineering, Linköping, Sweden, 1992.
- [3] V. Aue and J. H. Reed. "An interference robust CDMA demodulator that uses spectral correlation properties." In *Proceedings IEEE Vehicular Technology Conference*, pp. 563–567, IEEE, Piscataway, New Jersey, 1994.
- [4] P. A. Bello. "Characterization of time-variant linear channels." *IEEE Transactions on Communications Systems*, vol. 11, no. 4, pp. 360–393, December 1963.
- [5] R. E. Blahut. *Digital Transmission of Information*. Addison-Wesley, Reading, Massachusetts, 1990.
- [6] R. E. Blahut. *Principles and Practice of Information Theory*. Addison-Wesley, Reading, Massachusetts, 1990.
- [7] W. M. Brown and R. B. Crane. "Conjugate linear filtering." *IEEE Transactions on Information Theory*, vol. 15, no. 4, pp. 462–465, July 1969.
- [8] R. Cameron and B. D. Woerner. "An analysis of CDMA with imperfect power control." In *Proceedings IEEE Vehicular Technology Conference*, pp. 977–980, IEEE, Piscataway, New Jersey, 1992.
- [9] K. K. Chawla and D. C. Sarwate. "Parallel acquisition of PN sequences in DS/SS systems." *IEEE Transactions on Communications*, vol. 42, no. 5, pp. 2155–2164, May 1994.
- [10] R. C. Dixon. *Spread Spectrum Systems*. John Wiley & Sons, New York, New York, 2nd edition, 1984.

- [11] A. Duel-Hallen. "Performance of multiuser zero-forcing and MMSE decision-feedback detectors for CDMA channels." In *Proceedings IEEE Global Telecommunications Conference, Communication Theory Mini-Conference Volume*, pp. 82–86, IEEE, Piscataway, New Jersey, 1993.
- [12] K. Feher. *Advanced Digital Communications: Systems and Signal Processing Techniques*. Prentice-Hall, Englewood Cliffs, New Jersey, 1987.
- [13] G. H. Golub and C. F. van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, Maryland, 2nd edition, 1989.
- [14] H. Hashemi. "Simulation of the urban propagation channel." *IEEE Transactions on Vehicular Technology*, vol. 28, no. 3, pp. 213–225, August 1979.
- [15] S. Haykin. *Adaptive Filter Theory*. Prentice-Hall, Englewood Cliffs, New Jersey, 2nd edition, 1991.
- [16] J. M. Holtzman and L. M. A. Jalloul. "Rayleigh fading effect reduction with wideband DS/CDMA signals." In *Proceedings IEEE Global Telecommunications Conference*, pp. 553–557, IEEE, Piscataway, NJ, 1991.
- [17] IEEE Vehicular Technology Society Committee on Radio Propagation. "Received signal fading distribution." *IEEE Transactions on Vehicular Technology*, vol. 37, no. 1, pp. 57–60, February 1988.
- [18] R. A. Iltis. "A digital receiver for demodulation of CDMA waveforms with a-priori unknown delays and amplitudes." In *Proceedings IEEE Military Communications Conference*, pp. 113–116, IEEE, Piscataway, New Jersey, 1991.
- [19] W. C. Jakes, Jr. *Microwave Mobile Communications*. John Wiley & Sons, New York, New York, 1974.
- [20] E. A. Lee and D. G. Messerschmitt. *Digital Communication*. Kluwer Academic Publishers, Boston, Massachusetts, 1988.
- [21] W. C. Y. Lee. "Theory of wideband radio propagation." In *Proceedings IEEE Vehicular Technology Conference*, pp. 285–288, IEEE, Piscataway, NJ, 1991.
- [22] R. Lupas. *Near-Far Resistant Linear Multiuser Detection*. Ph.D. thesis, Princeton University, Dept. of Electrical Engineering, Princeton, New Jersey, January 1989.
- [23] R. Lupas and S. Verdú. "Linear multiuser detectors for synchronous code-division multiple-access channels." *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 123–136, January 1989.
- [24] R. Lupas and S. Verdú. "Near-far resistance of multiuser detectors in asynchronous channels." *IEEE Transactions on Communications*, vol. 38, no. 4, pp. 496–508, April 1990.

- [25] U. Madhow and M. L. Honig. "Error probability and near-far resistance of minimum mean square error interference suppression schemes for CDMA." In *Proceedings IEEE Global Telecommunications Conference*, pp. 1339–1343, IEEE, Piscataway, New Jersey, 1992.
- [26] The MathWorks, Inc. *MATLAB Reference Guide*. The MathWorks, Inc., Natick, Massachusetts, 1992.
- [27] S. L. Miller. "An adaptive direct-sequence code-division multiple-access receiver for multiuser interference rejection." Accepted for publication, *IEEE Transactions on Communications*.
- [28] T. K. Moon. *Parameter Estimation in SSMA Systems*. Ph.D. thesis, University of Utah, Dept. of Electrical Engineering, Salt Lake City, Utah, 1991.
- [29] T. K. Moon, R. T. Short, and C. K. Rushforth. "Average acquisition time for SSMA channels." In *Proceedings IEEE Military Communications Conference*, pp. 1042–1046, IEEE, Piscataway, New Jersey, 1991.
- [30] T. K. Moon, R. T. Short, and C. K. Rushforth. "A RASE approach to acquisition in SSMA systems." In *Proceedings IEEE Military Communications Conference*, pp. 1037–1041, IEEE, Piscataway, New Jersey, 1991.
- [31] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, New York, 3rd edition, 1991.
- [32] C. N. Pateros and G. J. Saulnier. "Interference suppression and multipath mitigation using an adaptive correlator direct sequence spread spectrum receiver." In *Proceedings IEEE International Conference on Communications*, pp. 662–666, IEEE, Piscataway, New Jersey, 1992.
- [33] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein. "Theory of spread-spectrum communications—A tutorial." *IEEE Transactions on Communications*, vol. 30, no. 5, pp. 855–884, May 1982.
- [34] H. V. Poor. "On parameter estimation in DS/CDMA formats." In W. A. Porter and S. C. Kak, editors, *Advances in Communications and Signal Processing*, pp. 59–70. Springer-Verlag, New York, New York, 1989.
- [35] J. C. Principe, B. de Vries, and P. G. de Oliveira. "The gamma filter—A new class of adaptive IIR filters with restricted feedback." *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 649–656, February 1993.
- [36] J. G. Proakis. *Digital Communications*. McGraw-Hill, New York, New York, 2nd edition, 1987.

- [37] A. Radović and B. Aazhang. "Iterative algorithms for joint data detection and delay estimation for code division multiple access communication systems." In *Proceedings of the 31st Annual Allerton Conference on Communication, Control, and Computing*, pp. 1–10, Allerton House, Monticello, Illinois, 1993.
- [38] S. S. Rappaport and D. M. Grieco. "Spread-spectrum signal acquisition: Methods and technology." *IEEE Communications Magazine*, vol. 22, no. 6, pp. 6–21, June 1984.
- [39] D. V. Sarwate and M. B. Pursley. "Crosscorrelation properties of pseudorandom and related sequences." *Proceedings of the IEEE*, vol. 68, no. 5, pp. 593–619, May 1980.
- [40] L. L. Scharf. *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*. Addison-Wesley, Reading, Massachusetts, 1991.
- [41] R. O. Schmidt. *A Signal Subspace Approach to Multiple Emitter Location and Spectral Estimation*. Ph.D. thesis, Stanford University, Stanford, California, November 1981.
- [42] R. A. Scholtz. "The spread spectrum concept." *IEEE Transactions on Communications*, vol. 25, no. 8, pp. 748–755, August 1977.
- [43] R. A. Scholtz. "The origins of spread-spectrum communications." *IEEE Transactions on Communications*, vol. 30, no. 5, pp. 822–854, May 1982.
- [44] M. K. Simon, R. A. Scholtz, B. K. Levitt, and J. K. Omura. *Spread Spectrum Communications*. Computer Science Press, Rockville, Maryland, 1985.
- [45] A. M. Slonneger and D. V. Sarwate. "Noncoherent parallel acquisition of PN sequences in direct-sequence spread-spectrum systems." In *IEEE Second International Symposium on Spread Spectrum Techniques and Applications*, pp. 31–34, IEEE, Piscataway, New Jersey, 1992.
- [46] P. Stoica and A. Nehorai. "MUSIC, maximum likelihood, and Cramér-Rao bound." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, pp. 720–741, May 1989.
- [47] E. G. Ström and S. L. Miller. "A reduced complexity adaptive near-far resistant receiver for DS-CDMA." In *Proceedings IEEE Global Telecommunications Conference*, pp. 1734–1738, IEEE, Piscataway, New Jersey, 1993.
- [48] E. G. Ström and S. L. Miller. "Optimum complexity reduction of minimum mean square error DS-CDMA receivers." In *Proceedings IEEE Vehicular Technology Conference*, pp. 568–572, IEEE, Piscataway, New Jersey, 1994.

- [49] E. G. Ström, S. Parkvall, S. L. Miller, and B. E. Ottersten. "Complexity reduction of direct-sequence code division multiple access receivers." In *Proceedings IEEE International Conference on Communications*, pp. 1643–1647, IEEE, Piscataway, New Jersey, 1994.
- [50] E. G. Ström, S. Parkvall, S. L. Miller, and B. E. Ottersten. "Propagation delay estimation of DS-CDMA signals in a fading environment." To appear in *Proceedings IEEE Global Telecommunications Conference, Communication Theory Mini-Conference Volume*, November 1994.
- [51] E. G. Ström, S. Parkvall, S. L. Miller, and B. E. Ottersten. "Sensitivity analysis of near-far resistant DS-CDMA receivers to propagation delay estimation errors." In *Proceedings IEEE Vehicular Technology Conference*, pp. 757–761, IEEE, Piscataway, New Jersey, 1994.
- [52] E. G. Ström, S. Parkvall, and B. E. Ottersten. "Near-far resistant propagation delay estimators for asynchronous direct-sequence code division multiple access systems." In C. G. Günther, editor, *Lecture Notes in Computer Science: Mobile Communications, Proceedings 1994 International Zurich Seminar on Digital Communications*, pp. 251–260. Springer-Verlag, New York, New York, March 1994.
- [53] H. Suzuki. "A statistical model for urban radio propagation." *IEEE Transactions on Communications*, vol. 25, no. 7, pp. 673–680, July 1977.
- [54] G. L. Turin. "Introduction to spread-spectrum antimultipath techniques and their application to urban digital radio." *Proceedings of the IEEE*, vol. 68, no. 3, pp. 328–352, March 1980.
- [55] H. L. van Trees. *Detection, Estimation and Modulation Theory*. John Wiley & Sons, New York, New York, 1968.
- [56] M. K. Varanasi and B. Aazhang. "Multistage detection in asynchronous code-division multiple-access communications." *IEEE Transactions on Communications*, vol. 38, no. 4, pp. 509–519, April 1990.
- [57] S. Verdú. "Minimum probability of error for asynchronous Gaussian multiple-access channels." *IEEE Transactions on Information Theory*, vol. 32, no. 1, pp. 85–96, January 1986.
- [58] A. J. Viterbi. "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm." *IEEE Transactions on Information Theory*, vol. 13, no. 2, pp. 260–269, April 1967.
- [59] R. B. Ward and K. P. Yiu. "Acquisition of pseudonoise signals by sequential estimation." *IEEE Transactions on Communications Technology*, vol. 13, no. 4, pp. 475–483, December 1965.

- [60] Z. Xie, C. K. Rushforth, R. T. Short, and T. K. Moon. "Joint signal detection and parameter estimation in multiuser communications." *IEEE Transactions on Communications*, vol. 41, no. 7, pp. 1208–1215, August 1993.
- [61] Z. Xie, R. T. Short, and C. K. Rushforth. "A family of suboptimum detectors for coherent multiuser communication." *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 4, pp. 683–690, May 1990.
- [62] S. Yoshida, A. Ushirokawa, S. Yanagi, and Y. Furuya. "DS/CDMA adaptive interference canceller on differential detection in fast fading channel." In *Proceedings IEEE Vehicular Technology Conference*, pp. 780–784, IEEE, Piscataway, New Jersey, 1994.
- [63] R. E. Ziemer and R. L. Peterson. *Digital Communications and Spread Spectrum Systems*. Macmillan, New York, New York, 1985.

BIOGRAPHICAL SKETCH

Karl Erik Gustav Ström was born June 28, 1965, in Örebro, Sweden. He completed his Master of Science in Electrical Engineering at the Royal Institute of Technology, Stockholm, Sweden, in February 1990. Ström entered the graduate program in electrical engineering at the University of Florida, Gainesville, in August 1990, and he is expected to graduate in December 1994 with a Doctor of Philosophy degree.

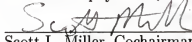
While still a student, Ström taught courses in circuits and electromagnetic field theory at the Royal Institute of Technology. Ström has since 1990 acted as a consultant for the Educational Group for Individual Development, Stockholm, Sweden.

Ström has accepted a postdoctoral position at the Royal Institute of Technology, where he will continue research in the DS-CDMA field.

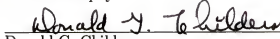
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Fred J. Taylor, Chairman,
Professor of Electrical Engineering

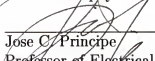
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Scott L. Miller, Cochairman,
Associate Professor of Electrical Engineering

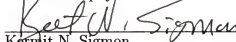
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Donald G. Childers
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

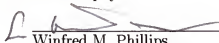

Jose C. Principe
Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Kermit N. Sigmon
Professor of Mathematics

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

December, 1994


Winfred M. Phillips
Dean, College of Engineering


Karen A. Holbrook
Dean, Graduate School